

Entwicklung und Anwendung eines fahrerorientierten Versuchskonzepts zur subjektiven Bewertung des Lenkgefühls am schweren Nutzfahrzeug

Am Fachbereich Maschinenbau
an der Technischen Universität Darmstadt



TECHNISCHE
UNIVERSITÄT
DARMSTADT

zur

Erlangung des Grades eines Doktor der Philosophie (Dr. phil.)
genehmigte

DISSERTATION

vorgelegt von

Alisa Stephanie Tanja Lindner, M.Sc. Psych., aus Coburg

Berichterstatter: Prof. Dr.-Ing. Ralph Bruder

Mitberichterstatter: Prof. Dr. Joachim Vogt

Tag der Einreichung: 24.01.2017

Tag der mündlichen Prüfung: 18.04.2017

Darmstadt 2017

D17

Danksagung

Zuerst möchte ich meinem Doktorvater Prof. Dr. Ralph Bruder danken. Er unterstützte mich in allen Phasen der Arbeit mit kritischen Diskussionen und wertvollen Anregungen. Dabei regte er mich immer wieder dazu an, meine Ergebnisse zu hinterfragen und die nächsten Schritte kritisch zu beleuchten. Eine Aussage von ihm ist mir besonders in Erinnerung geblieben: „Ich kann Ihnen nur die Flügel geben, fliegen müssen Sie selbst.“ Ich hoffe, dass mir dies hiermit gelungen ist.

Anschließend möchte ich meinem Zweitgutachter Prof. Dr. Joachim Vogt danken. Er begeisterte mich noch während meines Studiums für Ingenieurpsychologie und legte somit den Grundstein für meine berufliche Ausrichtung.

Mein großer Dank geht außerdem an Enrico Wohlfarth, meinen Betreuer bei der Daimler AG. Er gab mir die Chance, dieses interdisziplinäre Thema in der Lkw-Vorentwicklung zu bearbeiten. Dabei räumte er mir alle Freiheiten in der Ausrichtung meiner Forschung und der Durchführung der Studien ein.

Weiterhin möchte ich allen Mitarbeitern des IAD danken. Neben vielen anderen ist Dr.-Ing. Bettina Abendroth besonders hervorzuheben, die mir in allen Phasen der Arbeit ausgezeichnetes Feedback gab. Vielen Dank an alle, dass ihr mich als „Externe“ so gut aufnahmt.

Auch meine Kollegen von der Lkw-Vorentwicklung der Daimler AG haben zum Erfolg dieser Arbeit beigetragen. Mein besonderer Dank geht an Konrad Deuschle, der zu allen technischen Themen und darüber hinaus ein ausgezeichnete Ansprechpartner ist. Er nimmt sich immer die Zeit, meine „kleinen Fragen“ umfassend und fundiert zu beantworten. Weiterhin sind Thomas Wagner und Christopher Merkel zu nennen. Ohne ihren hervorragenden Beitrag bei technischen Aufgabenstellungen im Allgemeinen und der Fahrzeug-Betreuung im Speziellen wäre der Erfolg der Fahrversuche ausgeschlossen gewesen. Ferner danke ich Felix Feßler, Sabrina Heinrich, Kathrin Schmidt, Frieda Preuß, Karolin Striegel und Tanja Stoll für ihr Engagement bei der Durchführung ihrer studentischen Arbeiten. Letztere drei sind während unserer gemeinsamen Zeit in Zimmer 311 zu guten Freundinnen geworden. Mit euch allen als Kollegen wurden sogar wochenlange Fahrversuche nicht langweilig. Gut, dass wir uns getroffen haben.

Ein weiterer Dank geht an die Teilnehmer meiner Interviews und Fahrversuche. Dies waren zum einen Mitarbeiter verschiedener Bereiche der Daimler AG, zum anderen zahlreiche externe Lkw-Fahrer.

Ich danke außerdem Markus Heger von der Robert Bosch Automotive Steering GmbH. Er gab mir in der Schreibphase die nötigen Freiheiten, um die Arbeit fertig zu stellen.

Mein abschließender Dank geht an meine Familie, die mich jederzeit in meinen Vorhaben unterstützt. Allen voran ist hier meine Mutter Heike zu nennen, die das Fundament für meinen beruflichen Werdegang legte. Es ist gut zu wissen, dass sie mir immer zur Seite steht. Mein herzlicher Dank geht außerdem an meinen Mann Markus für seinen immerwährenden Rückhalt. Er ist immer für mich da, insbesondere wenn meine Ziele für mich gerade unerreichbar scheinen. Danke an euch für so vieles!



“Zwei Dinge sind zu unserer Arbeit nötig: Unermüdliche Ausdauer und die Bereitschaft, etwas, in das man viel Zeit und Arbeit gesteckt hat, wieder wegzuwerfen.”

Albert Einstein

Zusammenfassung

Die Lenkung ist ein elementarer Teil der Mensch-Maschine-Schnittstelle im Fahrzeug. Sie ist nicht nur essentiell für die Querführung des Fahrzeugs, sondern auch ein wichtiges Rückmeldeinstrument bezüglich des Fahrzustands. Die Summe der Wahrnehmungen eines Fahrers beim Lenken eines Fahrzeugs wird als Lenkgefühl bezeichnet. Bisher waren die Freiheiten bei der Lenkungsauslegung am schweren Nutzfahrzeug aufgrund der technischen Gegebenheiten stark eingeschränkt. Zukünftig werden jedoch hybride oder vollelektrische Lenksysteme an Bedeutung gewinnen. Mit ihnen kann die Lenkunterstützung besser an den Fahrerwunsch angepasst werden, was wiederum eine Optimierung des Lenkgefühls ermöglicht.

Es gibt sowohl praktische als auch wissenschaftliche Herangehensweisen, um das Lenkgefühl auf Basis der objektiven Kennwerte des Lenksystems zu beschreiben und zu optimieren. Die verwendeten Methoden sind jedoch oft auf Experten als Beurteiler ausgerichtet. Demgegenüber fordert der menschenzentrierte Gestaltungsprozess, dass der zukünftige Nutzer in den Produktentwicklungsprozess einbezogen wird. Die Nutzer, sog. Normalfahrer, haben im Gegensatz zu Experten keine besondere Expertise bezüglich der Querdynamik. Dementsprechend ist das Ziel dieser Arbeit die Entwicklung und Anwendung eines Versuchskonzepts, mit welchem Normalfahrer das Lenkgefühl am schweren Nutzfahrzeug bewerten können.

Zunächst werden Herangehensweisen zur subjektiven Bewertung und die gewonnenen Ergebnisse vorgestellt. Dabei wird herausgearbeitet, dass die Zusammensetzung der Stichprobe, der Versuchsaufbau und der Fragebogen die erzielten Ergebnisse maßgeblich beeinflussen. Nach Ableitung der Anforderungen wird das Versuchskonzept in einem iterativen Prozess unter Einbindung von Normalfahrern entwickelt. Das Konzept wird im Rahmen eines ersten Fahrversuchs evaluiert, in dem verschiedene Fahrzeuge im Straßenverkehr verglichen werden. Die Ergebnisse zeigen die Notwendigkeit auf, den Fragebogen zu optimieren und weitere Anforderungen festzulegen, insbesondere bzgl. der Stichprobe. Nachfolgend findet das Versuchskonzept in der zweiten Studie Anwendung. Der Vergleich verschiedener Fahrzeuge im Feld aber deckt die vermuteten Unterschiede im Lenkgefühl nicht auf. Die Analyse der Zuverlässigkeit der Bewertungen zeigt, dass Schwierigkeiten beim Vergleich verschiedener Fahrzeuge auftreten. Deshalb wird das Versuchskonzept hinsichtlich des Versuchsaufbaus weiter präzisiert. Die Variation des Lenkgefühls soll durch verschiedene Varianten innerhalb eines Fahrzeugs erfolgen. So sind auch unmittelbare Vergleiche zwischen Varianten sowie relative Bewertungen möglich.

Anschließend wird das Versuchskonzept dazu genutzt, Erkenntnisse zur Bewertung des Lenkgefühls zu generieren. Hierzu werden verschiedene Lenkmomentvarianten in synthetischen Manövern verglichen. Zum einen wird gezeigt, dass die Bewertungen der Fahrer zuverlässig sind und somit belastbare Ergebnisse vorliegen. Außerdem werden Ergebnisse zur Differenzierbarkeit der Varianten berichtet. Die Analysen zeigen ferner enge Zusammenhänge zwischen den subjektiven Bewertungen des Lenkgefühls und den objektiven Kennwerten auf. Es werden außerdem Erkenntnisse zum Zusammenhang zwischen den quasi-objektiven Einschätzungen und subjektiven Bewertungen generiert. Schlussendlich werden Empfehlungen zur Gestaltung der Lenkunterstützung in Abhängigkeit der Fahrsituation abgeleitet.

Abschließend wird weiterer Forschungsbedarf zur subjektiven Bewertung des Lenkgefühls vorgestellt.

Abstract

The steering system is an elementary part of the human-machine-interface in vehicles. On the one hand, it is essential for lateral guidance of a vehicle. On the other hand it is also an important source of feedback for the driver regarding the state of the vehicle. The sum of all the perceptions a driver has while steering a vehicle is called steering feel. Until now, the possibilities to tune steering systems in heavy trucks were limited due to technical properties. In the future, hybrid or electric steering systems will become increasingly important. In this way steering assist can be adapted to the drivers' needs more closely, enabling an optimisation of steering feel.

There are practical as well as empirical approaches to describe and optimise steering feel based upon objective parameters. However the most commonly used methods rely on assessment by technical experts. In contrast, the human-centered design process requires the inclusion of future users into the product development process. Unlike the technical experts, normal truck drivers do not have special expertise concerning lateral dynamics. Therefore, the goal of this work is the development and application of a test concept to enable normal drivers to evaluate steering feel in heavy trucks.

Firstly, approaches for subjective evaluation and achieved outcomes are presented. It is emphasized that the results achieved are determined by the composition of the sample, the experimental set-up and the questionnaire. After requirements are derived, the test concept is developed in an iterative process involving drivers. The concept is evaluated in a first driving study, in which different vehicles are compared in public traffic. The results highlight the need to improve the questionnaire and to determine further requirements, especially regarding the sample. Subsequently the test concept is applied in the second study. The comparison of different vehicles in the field does not uncover the assumed differences in steering feel. The analysis of the reliability of the ratings highlights difficulties when comparing different vehicles. Thus, the test concept is further specified in terms of the experimental set-up. The variation of steering feel should only be realized by implementing different configurations within one vehicle. In this way, more direct comparisons between configurations as well as relative ratings are possible.

In the last step the concept is used to generate findings about the subjective evaluation of steering feel. Different configurations of steering assist are compared in synthetic maneuvers. Firstly it is shown that the drivers' evaluations are stable and that the results achieved are reliable. Furthermore, results regarding the differentiability of the configurations are reported. Strong relations between the subjective evaluations and the objective parameters are found. Additionally, results concerning the relationship between the quasi-objective and the subjective evaluations are generated. Lastly, recommendations for the design of steering assist depending on the driving situation are given.

Finally, additional research was deemed necessary regarding the subjective evaluation of steering feel.

Abkürzungsverzeichnis

B	Bewertung
BV	Bewertungsvorgehen
FDR	false discovery rate
FS	Fahrsituation
F_HRST	Fahrer, der üblicherweise einen bestimmten Hersteller fährt
FZG	Fahrzeug
i	Invertierte Bewertungskriterien
KMO-Wert	Kaiser-Meyer-Olkin-Wert
NVH	Noise, Vibration und Harshness
Lkw	Lastkraftwagen
OP	Operationalisierung
Pkw	Personenkraftwagen
SP	Stichprobe
V	Variante
VD	Versuchsdesign

Formelverzeichnis

α	Wahrscheinlichkeit eines Typ I Fehlers / Cronbachs Index für interne Konsistenz
a_y	Querbeschleunigung [m/s ²]
B	Regressionskoeffizient
β	Wahrscheinlichkeit eines Typ II Fehlers / standardisierter Regressionskoeffizient
d	Cohens Maß der Effektstärke
df	Freiheitsgrade
δ	Lenkradwinkel [°]
F	Statistischer Kennwert mit bekannter Wahrscheinlichkeitsverteilung (F -Verteilung)
korrr. R^2	Anteil der durch das Modell erklärten Varianz an der gesamten Varianz bezogen auf die Grundgesamtheit
IQR	Interquartilabstand (inter quartil range)
M	Mittelwert
Mdn	Median
n	Fallzahl einer Sub-Stichprobe
N	Gesamt-Fallzahl
partielles η^2	Maß der Enge eines Zusammenhangs
p	Wahrscheinlichkeit
r	Pearsons Produkt-Moment Korrelations-Koeffizient
R^2	Anteil der durch das Modell erklärten Varianz an der gesamten Varianz
SD	Standardabweichung
t	Statistischer Kennwert mit bekannter Wahrscheinlichkeitsverteilung (t -Verteilung)
v	Fahrgeschwindigkeit [km/h]
X^2	Statistischer Kennwert mit bekannter Wahrscheinlichkeitsverteilung (X^2 -Verteilung)
z	Standardisierter Wert einer Statistik

Abbildungsverzeichnis

Abbildung 1. Hybrides Lenksystem mit hydraulischer und elektromechanischer (grün eingefärbt) Komponente (Quelle: Robert Bosch GmbH, 2016).	4
Abbildung 2. Regelkreis zur Erfüllung der Fahrzeugführungsaufgabe nach Harrer und Pfeffer (2013).	5
Abbildung 3. Das Lenkgefühl im engeren und weiteren Sinn nach Braess (2001; 2004).	6
Abbildung 4. Komponenten des Lenkgefühls.	7
Abbildung 5. Wesentliche Aspekte bei der Konzeption einer Studie zur subjektiven Bewertung des Lenkgefühls.	9
Abbildung 6. Absolute und relative Bewertung (B) am Beispiel drei bewerteter Varianten.	17
Abbildung 7. Skalenformate.	21
Abbildung 8. Schritte der Entwicklung des Fragebogens (gestrichelt) und der Definition der Anforderungen an den Versuchsaufbau (gepunktet) unter Angabe der verwendeten Quellen Experte (EXP), Fahrer (F), Literatur (L) und Normalfahrer (NF).	34
Abbildung 9. Skalen zur absoluten und relativen Niveau- und Gefallen-Bewertung.	40
Abbildung 10. Skalen zur absoluten und relativen Niveau- und Gefallen-Bewertung.	46
Abbildung 11. Versuchsstrecke mit beispielhaftem Streckenabschnitt zur Verdeutlichung des Bewertungsvorgehens (Kartenmaterial: Google Maps, 2014).	52
Abbildung 12. Die Mittelwerte der Niveau- und Gefallen-Bewertungen von <i>leichtgängig</i> und <i>Lenkungsstößigkeit</i> für die Sub-Stichproben F_HRST1 und F_HRST2 getrennt. Die Fehlerbalken repräsentieren die Standardabweichung.	55
Abbildung 13. Die Mittelwerte der Niveau- und Gefallen-Bewertungen von <i>agil</i> , <i>direkt</i> , <i>hat Spiel</i> und der Gesamt-Gefallenbewertung für die Sub-Stichproben F_HRST1 und F_HRST2 getrennt. Die Fehlerbalken repräsentieren die Standardabweichung.	57
Abbildung 14. Die Mittelwerte aller Niveau- und Gefallen-Bewertungen der beiden Fahrzeuge in Abhängigkeit der Sub-Stichproben F_HRST1 und F_HRST2. Die Fehlerbalken repräsentieren die Standardabweichung.	58
Abbildung 15. Die Mittelwerte aller Niveau- und Gefallen-Bewertungen der beiden Fahrzeuge in Abhängigkeit der Sub-Stichproben Normalfahrer und Experten. Die Fehlerbalken repräsentieren die Standardabweichung.	59
Abbildung 16. Die Mittelwerte der Effektstärke r und Standardabweichung SD in Abhängigkeit der Sub-Stichproben Normalfahrer und Experten. Die Fehlerbalken repräsentieren die Standardabweichung.	60
Abbildung 17. Die Mittelwerte der Gesamt-Gefallenbewertung der Sub-Stichproben Normalfahrer und Experten in Abhängigkeit der beiden Fahrzeuge. Die Fehlerbalken repräsentieren die Standardabweichung.	61
Abbildung 18. Vollständig benannte 7-stufige Skalen zur absoluten und relativen Niveau- und Gefallen-Bewertung.	73
Abbildung 19. Versuchsstrecke mit beispielhaftem Streckenabschnitt zur Verdeutlichung des Bewertungsvorgehens (Kartenmaterial: Google Maps, 2016).	78
Abbildung 20. Boxplots der Niveau- und Gefallen-Bewertungen von <i>Leichtgängig</i> , <i>Haltemoment</i> , <i>Rückstellverhalten-Geschwindigkeit</i> und <i>Rückstellverhalten-Restlenkwinkel</i> .	81
Abbildung 21. Boxplot der Niveau- und Gefallen-Bewertungen von <i>Lenkungsstößigkeit</i> .	81
Abbildung 22. Boxplots der Gesamt-Gefallenbewertung.	82

Abbildung 23. Mittelwerte der Niveau- und Gefallen-Bewertungen der Hin- und Rückfahrt. Die Fehlerbalken repräsentieren die Standardabweichung.	83
Abbildung 24. Mittelwerte der Niveau- und Gefallen-Bewertungen der regulären Fahrt und der Wiederholungsfahrt. Die Fehlerbalken repräsentieren die Standardabweichung.	83
Abbildung 25. Mittelwerte der Differenzen zwischen der regulären Fahrt und der Wiederholungsfahrt sowie zwischen den regulären Fahrten. Die Fehlerbalken repräsentieren die Standardabweichung.	84
Abbildung 26. Benannte 7-stufige Skalen zur relativen Niveau- und Gefallen-Bewertung.	96
Abbildung 27. Benannte 7-stufige Skalen zur relativen Niveau- und Gefallen-Bewertung.	98
Abbildung 28. Vom Fahrer aufzuwendende Lenkmomente der fünf Varianten.	105
Abbildung 29. Schematischer Ablauf der fünf Fahrmanöver.	107
Abbildung 30. Versuchsablauf am Beispiel des doppelten Spurwechsels.	110
Abbildung 31. Punktdiagramme und berechnete Regressionskurven der Gefallen-Bewertungen der acht Bewertungskriterien in Abhängigkeit der Niveau-Bewertungen. Die Sättigung der Punkte nimmt mit größer werdender Fallzahl zu.	117
Abbildung 32. Mittelwerte der Niveau-Bewertungen der fünf Varianten in den Kriterien <i>Lenkmoment (Rangieren)</i> , <i>Lenkmoment (Slalom)</i> , <i>Anlenkmoment</i> und <i>Haltemoment</i> . Die Fehlerbalken repräsentieren die Standardabweichung.	121
Abbildung 33. Punktdiagramme und berechnete Regressionskurven der Niveau-Bewertungen in den Kriterien <i>Lenkmoment (Rangieren)</i> , <i>Lenkmoment (Slalom)</i> , <i>Anlenkmoment</i> , <i>Haltemoment</i> und <i>Rückstellverhalten</i> in Abhängigkeit des maximalen Lenkmoments der Varianten. Die Sättigung der Punkte nimmt mit größer werdender Fallzahl zu.	122
Abbildung 34. Punktdiagramme und berechnete Regressionskurven der Gefallen-Bewertungen in den Kriterien <i>Lenkmoment (Rangieren)</i> , <i>Lenkmoment (Slalom)</i> , <i>Anlenkmoment</i> , <i>Haltemoment</i> und <i>Rückstellverhalten</i> in Abhängigkeit des maximalen Lenkmoments der Varianten. Die Sättigung der Punkte nimmt mit größer werdender Fallzahl zu.	123
Abbildung 35. Mittelwerte der Gefallen-Bewertungen der fünf Varianten in allen Bewertungskriterien. Die Fehlerbalken repräsentieren die Standardabweichung. Auf eine Darstellung der signifikanten Effekte wird aus Gründen der Übersichtlichkeit verzichtet.	126

Tabellenverzeichnis

Tabelle 1. Experten und Normalfahrer als Beurteiler des Lenkgefühls.	10
Tabelle 2. Die Vor- und Nachteile der Möglichkeiten zur Variation des Lenkgefühls unter Angabe der Folgen für die Gestaltung des Wechsels.	13
Tabelle 3. Die Vor- und Nachteile der zur Bewertung des Lenkgefühls geeigneten Versuchsumgebungen.	15
Tabelle 4. Zusammenstellung der Vor- und Nachteile sowie die mögliche Art der Bewertung in Abhängigkeit der Versuchsumgebung und der Variation des Lenkgefühls.	18
Tabelle 5. Überblick zu Bewertungskriterien ausgewählter Fragebögen unter Angabe der Zielgruppe.	19
Tabelle 6. Eigenschaften der Skalen ausgewählter Fragebögen unter Angabe der Zielgruppe.	22
Tabelle 7. Studien zum Lenkgefühl in Abhängigkeit der Versuchsumgebung, Variation des Lenkgefühls und Stichprobe*.	24
Tabelle 8. Überblick über die verwendete Methodik der Studien zum Lenkgefühl im Straßenverkehr.	25
Tabelle 9. Überblick über die verwendete Methodik der Studien zum Lenkgefühl auf dem Testgelände.	26
Tabelle 10. Überblick über die verwendete Methodik der Studien zum Lenkgefühl im Fahrsimulator.	29
Tabelle 11. Forschungsfragen und dazugehörige Forschungsarbeiten.	33
Tabelle 12. Auswahl der Bewertungskriterien mit dazugehöriger Erklärung.	39
Tabelle 13. Auswahl der Bewertungskriterien mit dazugehöriger Erklärung und Fahrsituation.	42
Tabelle 14. Erster Entwurf des Fragebogens.	44
Tabelle 15. Checkliste zum Versuchsaufbau.	47
Tabelle 16. Versuchsdesign unter Angabe der Stichprobengrößen.	49
Tabelle 17. Alter und Fahrpraxis der Normalfahrer.	50
Tabelle 18. Alter und Fahrpraxis der Experten.	50
Tabelle 19. Ergebnisse der Wilcoxon-Vorzeichen-Rang-Tests zum Vergleich der Bewertungen von <i>leichtgängig</i> und <i>Lenkungsstößigkeit</i> für die Sub-Stichproben F_HRST1 und F_HRST2 getrennt.	55
Tabelle 20. Ergebnisse der Wilcoxon-Vorzeichen-Rang-Tests zum Vergleich der Bewertungen von <i>agil</i> , <i>direkt</i> , <i>hat Spiel</i> und der Gesamt-Gefallenbewertung für die Sub-Stichproben F_HRST1 und F_HRST2 getrennt.	56
Tabelle 21. Ergebnisse der <i>t</i> -Tests für unabhängige Stichproben zum Vergleich aller Niveau- und Gefallen-Bewertungen der beiden Fahrzeuge in Abhängigkeit der Sub-Stichproben F_HRST1 und F_HRST2.	58
Tabelle 22. Antwortverhalten der Normalfahrer bei nicht-invertierten und invertierten Variablen.	62
Tabelle 23. Auswahl der Bewertungskriterien mit dazugehöriger Erklärung und Fahrsituation.	68
Tabelle 24. Versuchsdesign unter Angabe der Stichprobengrößen.	75
Tabelle 25. Ergebnisse von Friedmans ANOVAs bzw. Wilcoxon-Vorzeichen-Rang-Tests zum Vergleich der Bewertungen von <i>Leichtgängig</i> , <i>Haltemoment</i> , <i>Rückstellverhalten-Geschwindigkeit</i> und <i>Rückstellverhalten-Restlenkwinkel</i> .	80
Tabelle 26. Ergebnisse der Wilcoxon-Vorzeichen-Rang-Tests für die Kriterien, deren Gefallen-Bewertungen der regulären Fahrt und Wiederholungsfahrt vor der α -Fehler-Korrektur signifikant verschieden waren.	84
Tabelle 27. Zweifaktorielle Lösung unter Angabe der internen Konsistenz, Faktorladung und Trennschärfe.	86
Tabelle 28. Antwortverhalten der Probanden bei nicht-invertierten und invertierten Variablen.	88

Tabelle 29. Zusammenstellung der Vor- und Nachteile sowie die mögliche Art der Bewertung in Abhängigkeit der Versuchsumgebung und der Variation des Lenkgefühls.	92
Tabelle 30. Bewertungskriterien mit dazugehöriger Erklärung und Fahrsituation.	94
Tabelle 31. Bewertungskriterien mit dazugehöriger Erklärung und Fahrsituation.	97
Tabelle 32. Checkliste zum Versuchsaufbau.	99
Tabelle 33. Versuchsdesign unter Angabe der Stichprobengrößen.	103
Tabelle 34. Beschreibung der Fahrmanöver unter Angabe der Fahrgeschwindigkeit v , des maximalen Lenkradwinkels δ und der maximalen Querbesehleunigung a_y .	106
Tabelle 35. Bewertungskriterien mit dazugehöriger Erklärung und Fahrsituation.	109
Tabelle 36. Zweifaktorielle Lösung unter Angabe der internen Konsistenz, Faktorladung und Trennschärfe.	113
Tabelle 37. Antwortverhalten der Probanden.	114
Tabelle 38. Zusammenfassung der Regressionsmodelle zur Vorhersage der Niveau-Bewertung der acht Bewertungskriterien auf Basis der Gefallen-Bewertungen.	116
Tabelle 39. Prozentuale Häufigkeiten der korrekt wahrgenommenen, nicht wahrgenommenen und inkorrekt wahrgenommenen Unterschiede zwischen den fünf Varianten für die Kriterien <i>Lenkmoment (Rangieren)</i> , <i>Lenkmoment (Slalom)</i> , <i>Anlenkmoment</i> und <i>Haltemoment</i> .	119
Tabelle 40. Ergebnisse der Varianzanalysen mit Messwiederholung zum Vergleich der Niveau-Bewertungen der fünf Varianten in den Kriterien <i>Lenkmoment (Rangieren)</i> , <i>Lenkmoment (Slalom)</i> , <i>Anlenkmoment</i> und <i>Haltemoment</i> .	120
Tabelle 41. Zusammenfassung der Regressionsmodelle zur Vorhersage der Niveau-Bewertung in den Kriterien <i>Lenkmoment (Rangieren)</i> , <i>Lenkmoment (Slalom)</i> , <i>Anlenkmoment</i> , <i>Haltemoment</i> und <i>Rückstellverhalten</i> auf Basis des maximalen Lenkmoments der Varianten.	122
Tabelle 42. Zusammenfassung der Regressionsmodelle zur Vorhersage der Gefallen-Bewertung in den Kriterien <i>Lenkmoment (Rangieren)</i> , <i>Lenkmoment (Slalom)</i> , <i>Anlenkmoment</i> , <i>Haltemoment</i> und <i>Rückstellverhalten</i> auf Basis des maximalen Lenkmoments der Varianten.	123
Tabelle 43. Ergebnisse der Varianzanalysen mit Messwiederholung zum Vergleich der Gefallen-Bewertungen der fünf Varianten in allen Bewertungskriterien.	125
Tabelle 44. Gestaltungsempfehlungen in Abhängigkeit der Fahrsituation.	126
Tabelle 45. Forschungsfragen, dazugehörige Forschungsarbeiten und Ergebnisse.	141

Inhaltsverzeichnis

1. Einleitung	1
1.1. Motivation	1
1.2. Ziel und Struktur der Arbeit	2
2. Stand der Technik und Forschung	4
2.1. Stand der Technik	4
2.2. Stand der Forschung zum Begriff des Lenkgefühls	5
2.3. Stand der Forschung zur Bewertung des Lenkgefühls	8
2.3.1. Stichprobe	9
2.3.2. Versuchsaufbau	12
2.3.3. Fragebogen	18
2.4. Erkenntnisse zur Bewertung des Lenkgefühls	23
2.4.1. Studien im Straßenverkehr	24
2.4.2. Studien auf dem Testgelände	25
2.4.3. Studien im Fahrsimulator	28
2.5. Ableitung des Forschungsbedarfs	29
3. Entwicklung und Evaluation eines Versuchskonzepts zur subjektiven Bewertung des Lenkgefühls durch Normalfahrer	34
3.1. Entwicklung des Versuchskonzepts	34
3.1.1. Entwicklung des Fragebogens	34
3.1.2. Versuchsaufbau	41
3.1.3. Entwurf des Versuchskonzepts	43
3.2. Erste Evaluation des Versuchskonzepts: Bewertung des Lenkgefühls verschiedener Fahrzeuge (Studie I)	47
3.2.1. Ziele und Hypothesen	47
3.2.2. Methodik	48
3.2.3. Ergebnisse	54
3.2.4. Diskussion	63
3.2.5. Erste Optimierung des Versuchskonzepts	66
3.3. Zweite Evaluation des Versuchskonzepts: Bewertung des Lenkgefühls verschiedener Fahrzeuge (Studie II)	74
3.3.1. Ziele und Hypothesen	74
3.3.2. Methodik	75
3.3.3. Ergebnisse	79
3.3.4. Diskussion	88
3.3.5. Zweite Optimierung des Versuchskonzepts	91
3.4. Versuchskonzept	96
3.4.1. Fragebogen	96

3.4.2. Gesammelte Anforderungen an den Versuchsaufbau	98
4. Anwendung des Versuchskonzepts: Bewertung des Lenkgefühls verschiedener Varianten der Lenkunterstützung (Studie III)	100
4.1. Ziele & Hypothesen	100
4.2. Methodik	102
4.2.1. Versuchsdesign	102
4.2.2. Stichprobe	103
4.2.3. Variation des Lenkgefühls	103
4.2.4. Versuchsumgebung und Festlegen relevanter Fahrsituationen	105
4.2.5. Art der Bewertung, Bewertungskriterien und Bewertungsskala	107
4.2.6. Versuchsablauf	109
4.2.7. Auswertungsvorgehen	111
4.3. Ergebnisse	112
4.3.1. Summative Evaluation des Fragebogens	112
4.3.2. Vorhersagbarkeit der Gefallen-Bewertungen in Abhängigkeit der Niveau-Bewertung	116
4.3.3. Zuverlässigkeit der Bewertung	117
4.3.4. Wahrnehmbarkeit der Unterschiede zwischen den Varianten	118
4.3.5. Vorhersagbarkeit der Niveau-Bewertung in Abhängigkeit der Variante	121
4.3.6. Vorhersagbarkeit der Gefallen-Bewertungen in Abhängigkeit der Variante	122
4.3.7. Explorative Analysen	123
4.3.8. Ableitung der Gestaltungsempfehlungen	124
4.4. Diskussion	127
4.4.1. Summative Evaluation des Versuchskonzepts	127
4.4.2. Vorhersagbarkeit der Gefallen-Bewertungen in Abhängigkeit der Niveau-Bewertung	129
4.4.3. Zuverlässigkeit der Bewertung	131
4.4.4. Unterschiedliche Niveau-Bewertung der Varianten	133
4.4.5. Vorhersagbarkeit der Niveau-Bewertung in Abhängigkeit der Variante	134
4.4.6. Vorhersagbarkeit der Gefallen-Bewertungen in Abhängigkeit der Variante	135
4.4.7. Explorative Analysen	136
4.4.8. Ableitung der Gestaltungsempfehlungen	137
4.4.9. Limitationen des Versuchs	138
5. Beantwortung der Forschungsfragen	141
6. Ausblick	144
7. Literaturverzeichnis	146
8. Anhang	156

1. Einleitung

1.1. Motivation

Die Querführung eines Kraftfahrzeugs erfolgt gewöhnlich mithilfe des Lenkrads und bildet zusammen mit der Längsführung über die Pedalerie die primäre Fahraufgabe (Bubb, 2002). Die Lenkung stellt somit einen elementaren Teil der Fahrzeugführung dar und beeinflusst den Fahrkomfort maßgeblich. Während des Fahrens nutzt der Fahrer das Lenkrad, um die gewünschten Richtungsänderungen an das Fahrzeug weiterzugeben. Gleichzeitig gibt das Lenkrad auch Rückmeldungen an den Fahrer, beispielsweise bei Fahrbahnunebenheiten. Dieses Zusammenspiel aus Lenkwinkleingaben, Feedback und Fahrzeugreaktion wird als Lenkgefühl beschrieben (Harrer & Pfeffer, 2013).

Aufgrund der Bedeutung für die Fahrzeugführung muss das Lenksystem hinsichtlich Sicherheit (Heiβing & Ersoy, 2011) und Ergonomie (Breuer, 2000) optimal gestaltet werden. Insbesondere am Nutzfahrzeug ist es von großer Relevanz, die Lenkung optimal an die Bedürfnisse des Fahrers anzupassen. Erstens sind die aufzubringenden Lenkmomente aufgrund der Masse des Lastkraftwagen (Lkw) deutlich größer als am Personenkraftwagen (Pkw), weshalb eine Unterstützung durch Hilfskraftlenkungen besonders im niedrigen Geschwindigkeitsbereich erforderlich ist. Ferner ist die Querführungsaufgabe im Lkw in vielen Fahrsituationen anspruchsvoller als im Pkw, beispielsweise aufgrund der größeren Fahrzeugbreite. Darüber hinaus werden Lastkraftwagen zu einem Großteil gewerblich genutzt, sodass der Fahrersitz gleichzeitig einen elementaren Teil des Arbeitsplatzes eines Berufskraftfahrers darstellt. Deshalb sollte die Lenkung möglichst gut an die Fahrerbedürfnisse angepasst werden.

Die technischen Fortschritte im Bereich der Lenkungsentwicklung machen jedoch immer mehr Anpassungen möglich (Schöttle, 2014). Im Pkw-Bereich hat sich der Wandel von der hydraulischen zur elektrischen Servolenkung bereits vollzogen. Im Lkw werden heute hybride Lenksysteme eingesetzt, welche die klassischen hydraulischen Servolenkungen um eine elektromechanische Komponente ergänzen (Wiesel, 2010). So beschränkt sich die Gestaltung des Lenkgefühls nicht mehr nur auf einen Kompromiss zwischen dem Entwicklungsziel einer komfortablen Lenkbarkeit bei niedrigen Fahrgeschwindigkeiten und angenehmer Spurführung bei hohen Fahrgeschwindigkeiten (Heiβing & Brandl, 2002). Vielmehr ist nun eine geschwindigkeitsabhängige Lenkmomentgestaltung umsetzbar, wie sie im Pkw bereits Stand der Technik ist. Zusätzlich ist ein aktiver Rücklauf des Lenkrads realisierbar.

Aus den frei gestaltbaren Kennlinien der elektrischen Systeme ergibt sich eine Fülle an Auslegungsmöglichkeiten. Diese bringt die Notwendigkeit mit sich, die Zusammensetzung des Lenkgefühls zu verstehen und dieses anschließend zu optimieren. In der Praxis wird die Lenkung heute im Wesentlichen auf Basis von Abstimmungsfahrten ausgelegt (Schimmel, 2010). Dabei wird jedoch wenig Augenmerk darauf gelegt, inwiefern die Kundenmeinung mit diesem Prozess wirklich abgebildet wird (Pfeffer & Scholz, 2010).

Es gibt jedoch auch wissenschaftliche Herangehensweisen, welche ein tieferes Verständnis des Lenkgefühls als Ziel verfolgen (Zschocke, 2009). Dazu werden größtenteils standardisierte Fahrversuche durchgeführt. In diesen vergleichen Fahrer das aus verschiedenen Lenkungsauslegungen resultierende Lenkgefühl. Die Variation des Lenkgefühls kann entweder durch den Einsatz verschiedener Fahrzeuge oder durch die Realisierung

verschiedener Parametrierungen innerhalb eines Fahrzeugs erfolgen. Das Lenkgefühl wird dann von Fahrern in entsprechenden Fahrsituationen beobachtet und anschließend mithilfe von Fragebögen oder Checklisten bewertet und dokumentiert (Ackert, 2008). Anschließend können die gewonnenen subjektiven Bewertungen in Zusammenhang mit den vorher gemessenen objektiven Parametern gesetzt werden (Neukum, Krüger & Schuller, 2001). So kann die für die jeweilige Situation geeignetste Parametrierung identifiziert und schlussendlich im Fahrzeug eingesetzt werden. Darüber hinaus wird so eine Datengrundlage für die simulative Auslegung gewonnen, welche immer mehr an Bedeutung gewinnt.

Als Beurteiler¹ werden hierfür meist Experten aus dem Bereich der Querdynamik eingesetzt (Koch, 2010; Neukum et al., 2001), da diese besondere Kenntnisse und Fähigkeiten bei der Beurteilung der Lenkung vorweisen können. Nach dem menschenzentrierten Gestaltungsprozess (ISO 9241-210:2010) sollen jedoch die zukünftigen Nutzer eines Produkts in den gesamten Produktentwicklungsprozess eingeschlossen werden. Dies bedeutet, dass Normalfahrer in den Prozess der Lenkungsauslegung eingebunden werden müssen. Normalfahrer zeichnen sich dadurch aus, dass sie keine besonderen Fertigkeiten bezüglich des querdynamischen Grenzbereichs vorweisen können (Ullmann, 2006). Nach Badiru (2014) liegen zahlreiche Erkenntnisse zu den Vorlieben von Kunden im Bereich des allgemeinen Fahrverhaltens vor, nicht aber spezifisch für das Lenkverhalten. Bezüglich des Lenkgefühls aus Expertensicht sind vor allem die Arbeiten von Zschocke (2009) und Harrer (2007) zu nennen. Die Breite und Tiefe der Erkenntnisse aus Fahrersicht sind im Vergleich dazu geringer, wobei beispielsweise bei Anand (2014) und Decker (2009) Ergebnisse zu finden sind. Ein noch größeres Forschungsdefizit besteht jedoch in Bezug auf das Lenkgefühl im Lkw, da hier lediglich einzelne Arbeiten vorliegen (Rothhämel, 2013). Diese Forschungslücken lassen sich möglicherweise darauf zurückführen, dass bei der Nutzung von Normalfahrern als Beurteiler ein für sie geeignetes Versuchskonzept nötig ist. Dies beinhaltet zum einen, dass der verwendete Fragebogen der Terminologie und Expertise der Fahrer entsprechen muss (Kallus, 2010). Zum anderen muss das gesamte Versuchskonzept auf die Erfahrung und das fahrerische Können von Normalfahrern ausgerichtet sein. Nur so können mit ihnen zuverlässige Ergebnisse erzielt werden.

1.2. Ziel und Struktur der Arbeit

In Anlehnung an den menschenzentrierten Gestaltungsprozess (ISO 9241-210:2010) ergibt sich somit die Notwendigkeit, die von Experten zur Bewertung des Lenkgefühls im Pkw entwickelten Methoden anzupassen. Ziel dieser Arbeit ist folglich die Entwicklung eines geeigneten Versuchskonzepts, mit dem das Lenkgefühl am schweren Nutzfahrzeug durch Normalfahrer bewertet werden kann. Dieses soll im Anschluss dazu genutzt werden, Erkenntnisse über das Lenkgefühl zu generieren.

¹ Aus Gründen der besseren Lesbarkeit wird im Text verallgemeinernd das generische Maskulinum verwendet. Diese Formulierungen umfassen gleichermaßen weibliche und männliche Personen.

In Kapitel 1.1 wurde das Themengebiet der subjektiven Bewertung des Lenkgefühls vorgestellt. Ferner wurde die bestehende Notwendigkeit, den Fahrer in die Lenkungsauslegung am Lkw einzubeziehen, hervorgehoben. In diesem Kapitel werden Ziele und Struktur der Arbeit dargelegt.

Inhalt von Kapitel 2 ist ein Überblick zum Stand der Forschung und Technik. Hierbei werden aktuelle Entwicklungen auf dem Bereich der Lenksysteme im Lkw vorgestellt (Kapitel 2.1). Nach einer Definition des Begriffs Lenkgefühl (Kapitel 2.2) werden anschließend relevante Aspekte bei der subjektiven Bewertung des Lenkgefühls vorgestellt (Kapitel 2.3). Dabei liegt der Fokus auf der Stichprobe, dem Versuchskonzept und dem Fragebogen. Schlussendlich werden Ergebnisse aus vorherigen Studien zum Lenkgefühl in Abhängigkeit der Versuchsumgebung zusammengefasst (Kapitel 2.4). Auf Basis dieser Zusammenstellung werden die Forschungsfragen der vorliegenden Arbeit abgeleitet (Kapitel 2.5).

Im dritten Kapitel wird, aufbauend auf die abgeleiteten Anforderungen an das Versuchskonzept, die iterative Entwicklung des Fragebogens vorgestellt (Kapitel 3.1). Im Anschluss werden die Ergebnisse der Evaluation im Rahmen von Studie I berichtet, bei der das Lenkgefühl zweier Lkw im Straßenverkehr bewertet wurde (Kapitel 3.2). Nach der Darstellung der weiteren Optimierung des Konzepts folgt die Vorstellung von Studie II (Kapitel 3.3). Diese hatte zum Ziel, das Lenkgefühl mehrerer Lkw im Straßenverkehr zu evaluieren. Den Abschluss dieses Kapitels bildet die erneute Optimierung des Versuchskonzepts.

Das vierte Kapitel schildert die Anwendung dieses Versuchskonzepts in einem Fahrversuch zum Vergleich verschiedener Varianten des Lenkgefühls in synthetischen Manövern. Nach dem Ableiten der Hypothesen (Kapitel 4.1) und der Vorstellung der Methoden (Kapitel 4.2) werden die Ergebnisse hinsichtlich der aufgedeckten Zusammenhänge zwischen den objektiven Parametern und subjektiven Bewertungen präsentiert (Kapitel 4.3). Außerdem werden Analysen zur Zuverlässigkeit der Bewertungen präsentiert. Eine Diskussion der Ergebnisse bildet den Abschluss dieses Kapitels (Kapitel 4.4).

Die Ergebnisse der Forschungsarbeiten werden in Kapitel 5 zur Beantwortung der Forschungsfragen herangezogen. Den Abschluss dieser Arbeit bildet ein Ausblick (Kapitel 6).

2. Stand der Technik und Forschung

2.1. Stand der Technik

In den ersten Automobilen waren rein mechanische Lenksysteme verbaut. Mit fortschreitender Weiterentwicklung aller Komponenten des Fahrzeugs hielten hydraulische Servolenkungen Einzug (Karch & Grüner, 2007). Diese erwirkten im Vergleich zu manuellen Lenksystemen eine Reduktion der notwendigen Lenkmomente und damit eine Steigerung des Fahrkomforts. Bei der Auslegung war es jedoch stets nötig, einen Kompromiss zwischen ausreichend niedrigen Parkiermomenten und ausreichend hohen Lenkmomenten bei schneller Fahrt zu finden (Ackert, 2008). Heute werden im Pkw hauptsächlich elektrische Servolenkungen verwendet, bei welchen ein Elektromotor das vom Fahrer aufgewendete Lenkmoment verstärkt. Diese Systeme bieten neben einer Kraftstoffersparnis (Müller, 2004) den Vorteil, dass die Lenkunterstützung noch besser an den Fahrerwunsch angepasst werden kann (Barthenheier, 2004). Außerdem ermöglichen sie die Umsetzung von Fahrerassistenzsystemen, beispielsweise Querführungs- oder Stau-Assistenten. Mit dem Einsatz von Elektrolenkungen im Lkw könnten diese Vorteile auch dort realisiert werden.

Bisher lassen sich solche Systeme jedoch nicht mit vertretbarem technischen Aufwand realisieren (Tagesson, 2014). Deshalb werden bis auf wenige Ausnahmen immer noch Kugelumlauf Lenkungen mit hydraulischer Servolenkung eingesetzt. Es sind jedoch bereits einige hybride Lenksysteme auf dem Markt (Abbildung 1).

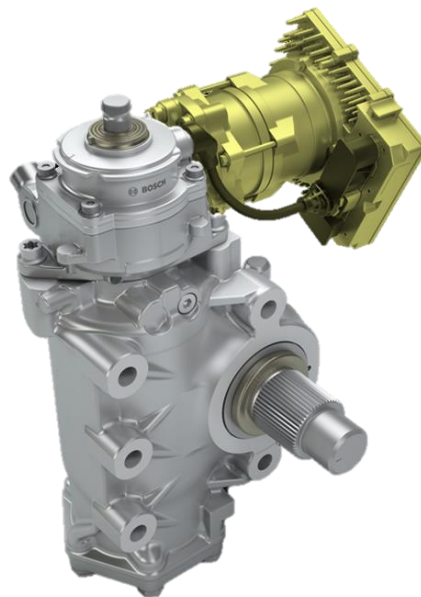


Abbildung 1. Hybrides Lenksystem mit hydraulischer und elektromechanischer (grün eingefärbt) Komponente (Quelle: Robert Bosch GmbH, 2016).

Diese zeichnen sich dadurch aus, dass die klassische hydraulische Servolenkung mit einer elektronischen Antriebs- und Steuereinheit gekoppelt ist (Robert Bosch GmbH, 2016). Aus Fahrersicht sind die folgenden drei Verbesserungen von besonderer Bedeutung:

Erstens kann durch den Elektromotor, wie beim Pkw, das vom Fahrer aufzubringende Lenkmoment geschwindigkeitsabhängig gestaltet werden (Weinmann, Bitzer, Boos & Burkhart, 2015). Somit ist das Unterstützungsmoment im Stand bzw. bei niedrigen Geschwindigkeiten deutlich höher, womit eine Reduktion des vom Fahrer aufzubringenden Lenkmoments einhergeht. Dieses Unterstützungsmoment fällt bei hohen Geschwindigkeiten niedriger aus, wodurch eine höhere Beherrschbarkeit erzielt wird. Zum zweiten kann der aktive Rücklauf, also das Zurückdrehen des Lenkrads aus dem eingelenkten Zustand in die Geradeausstellung, variiert werden (Robert Bosch GmbH, 2016). Dieses Rückstellmoment kann ebenfalls stärker oder schwächer gestaltet werden. Damit einher geht eine Modifizierung des Restlenkwinkels, der vom Fahrer selbst zurückgelenkt werden muss. Hierbei ist ebenfalls eine geschwindigkeitsabhängige Auslegung möglich. Darüber hinaus kann diese sogenannte aktive Lenkung auch ohne ein vom Fahrer aufgebrachtes Lenkmoment eingreifen. Somit kann der Fahrer durch Fahrerassistenzsysteme bei der Querführung unterstützt werden (Wiesel, Schwarzhaupt, Frey & Gauterin, 2010). Diese Unterstützung kann von einem Lane-Keep-Assist bis hin zum hochautomatisierten Fahren reichen.

2.2. Stand der Forschung zum Begriff des Lenkgefühls

Entsprechend des Regelkreises zur Erfüllung der Fahrzeugführungsaufgabe nach Harrer und Pfeffer (2013) kann die Interaktion zwischen Fahrer und Fahrzeug mit vier Komponenten beschrieben werden (Abbildung 2).

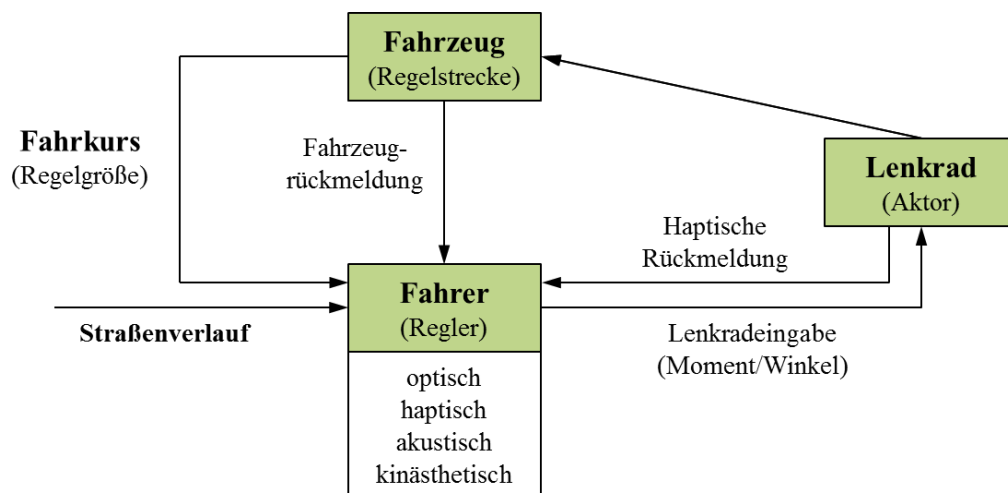


Abbildung 2. Regelkreis zur Erfüllung der Fahrzeugführungsaufgabe nach Harrer und Pfeffer (2013).

Der Fahrer (Regler) beeinflusst die physikalischen Größen so, dass ein vorgegebener Fahrkurs (Soll-Wert der Regelgröße) möglichst gut eingehalten wird. Das Lenkrad (Aktor) als Fahrer-Fahrzeug-Schnittstelle ist das Stellglied, das Signale einer Regelung an das Fahrzeug (Regelstrecke) weiterleitet. Dieses ist die zu regelnde physikalische Größe, auf die der Regler über den Aktor einwirkt. Die Regelgröße ist die durch Regelung konstant oder gezielt veränderlich zu haltende Größe. Aus diesem Zusammenspiel entsteht das Lenkgefühl, welches wie folgt definiert wird (Harrer & Pfeffer, 2013, S. 127):

„Lenkgefühl ist die Summe der optischen, kinästhetischen und haptischen Sinneseindrücke des Fahrers beim Lenken eines Fahrzeugs und entspricht einer subjektiv empfundenen, komplexen Erfahrung.“

Zum optischen Feedback gehört beispielweise die visuell wahrgenommene Abweichung vom gewünschten Kurs (Harrer, 2007). Nach Tomaske und Meywerk (2006) wird der Lenkwinkel ebenfalls hauptsächlich mit dem Auge wahrgenommen. Unter die kinästhetischen Wahrnehmungen fallen Faktoren der Fahrzeugreaktion (Harrer & Pfeffer, 2013), beispielsweise die Querschleunigung. Nach Decker (2009) ist an der Wahrnehmung fahrdynamischer Größen jedoch der vestibuläre Sinneskanal beteiligt. Die für das Lenkgefühl wesentlichen Aspekte werden jedoch haptisch wahrgenommen. Hierunter fallen die am Lenkrad anliegenden Lenkmomente und nach Harrer und Pfeffer (2013) auch die Lenkwinkel.

Nach Bubb und Schmidtke (1993) ist die Bedeutung der einzelnen Sinneskanäle für die Fahrzeugführung von der Fahrerfahrung abhängig. So orientieren sich erfahrene Fahrer weniger an der optisch wahrnehmbaren Abweichung vom Sollkurs, sondern vielmehr an der kinästhetisch wahrgenommenen Fahrtrichtung.

Der Fahrer bringt diese Informationen mit seinen Erwartungen und bisherigen Erfahrungen in Bezug und subjektiviert diese Wahrnehmungen. Hieraus resultiert, dass das Lenkgefühl nicht vollständig mit Kennwerten und Gleichungen erklärt werden kann (Gil Gómez, Nybacka, Bakker & Drugge, 2015). Auch Rothhämel, Ijkema und Drugge (2010) heben die Subjektivität des Lenkgefühls hervor, indem sie es mit dem Gewicht eines Gegenstands vergleichen. So ist die subjektive Einschätzung eines Sachverhalts von vielen Faktoren abhängig. Zum einen können inter-individuelle Faktoren wie beispielsweise die körperliche Statur eine Rolle spielen. Zum anderen sind intra-individuelle Faktoren wie die Tagesform mögliche Einflussfaktoren.

Es wird zwischen dem Lenkgefühl im engeren und im weiteren Sinn unterscheiden (Braess, 2001, 2004; Harrer, 2007; Wolf, 2009) (Abbildung 3).

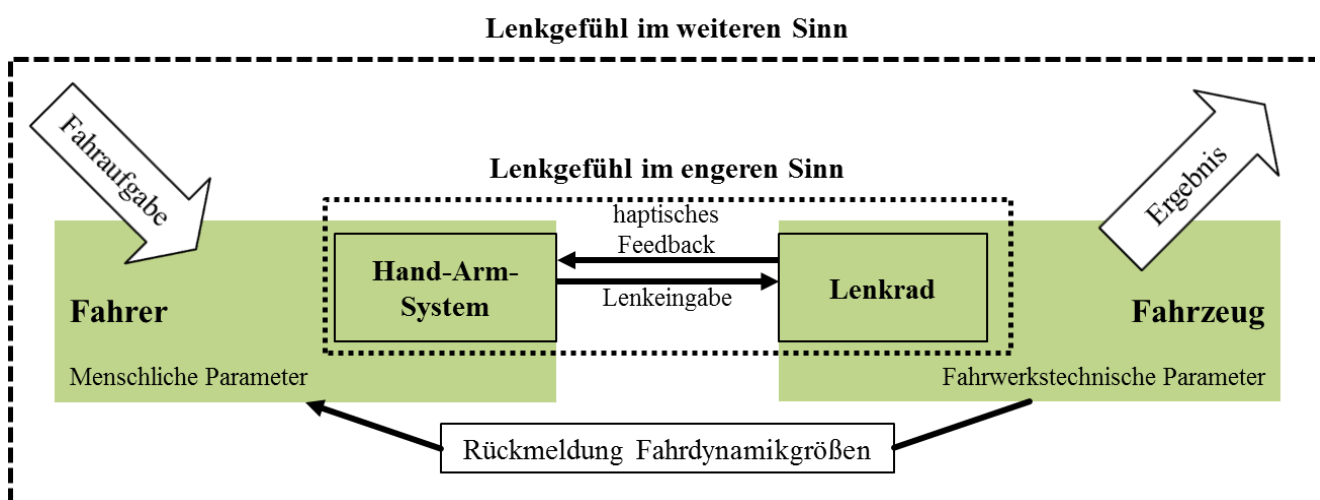


Abbildung 3. Das Lenkgefühl im engeren und weiteren Sinn nach Braess (2001; 2004).

Das Lenkgefühl im engeren Sinn umfasst lediglich die Interaktion zwischen dem Hand-Arm-System und dem Lenkrad sowie die haptisch wahrgenommene Rückmeldung (gepunktete Linie). Zusätzlich zur unmittelbaren

Interaktion am Lenkrad schließt das Lenkgefühl im weiteren Sinne auch Fahrzeugreaktionen mit ein (gestrichelte Linie). Hierzu zählen beispielsweise die Querbeschleunigung sowie die Wank- und Gierreaktion. Anstelle der Haptik sind hier die optische und kinästhetische Wahrnehmung von Bedeutung (Wolf, 2009). Das Lenkgefühl im weiteren Sinne entspricht somit der Definition von Harrer und Pfeffer (2013).

Zusätzlich zu den oben vorgestellten Definitionen des Lenkgefühls versuchen viele Autoren, das Lenkgefühl in die zugrunde liegenden Faktoren aufzugliedern. Darunter ist zu verstehen, dass die subjektiven Wahrnehmungen des Fahrers thematisch untergliedert werden. Mit diesem Ansatz wird versucht, die Zusammensetzung des Lenkgefühls genauer zu beschreiben. Somit sind sie auch mit den Skalen eines Fragebogens vergleichbar.

Beispielsweise bei Buschardt (2003) liegt der Fokus auf den objektiven Faktoren, welche dem Lenkgefühl zugrunde liegen. Im Gegensatz dazu stellt Wolf (2009) insgesamt 13 fahrer- und fahrzeugabhängige Faktoren dar, welche unter anderem die Fahraufgabe, aber auch die Lenkradgeometrie und Fahrerphysiologie beinhalten. Solche Klassifizierungen wurden für die folgende Übersicht nicht berücksichtigt, da sie nur begrenzt Aussagen zu den zugrundeliegenden Kategorien der subjektiven Wahrnehmungen beim Lenken ermöglichen. Im Folgenden wird ein Überblick über in der Literatur zu findende Kategorien des Lenkgefühls gegeben (Barthenheier, 2004; Braess, 2001; Braess & Seiffert, 2013; Farrer, 1993; Harrer & Pfeffer, 2013; Heißing & Brandl, 2002; Heißing & Ersoy, 2011; Koide & Kawakami, 1988; Mandhata et al., 2012; Pietsch & Heißing, 2010; Resch & Mast, 2006; Rothhämel, 2013):

Es zeigt sich, dass die Kategorien sowohl in ihrer Anzahl als auch in ihrem Inhalt stark variieren. Bei Barthenheier (2004) beispielsweise werden nur drei Kategorien zur Untergliederung genutzt, nämlich Komfort, Sportlichkeit und empfundene Sicherheit. Harrer und Pfeffer (2013) formulieren die vier Oberziele Lenkpräzision, Lenkkomfort, Lenkungsrückmeldung und Lenkdynamik. Rothhämel et al. (2010) hingegen extrahierten aus ihrem entwickelten Fragebogen durch multidimensionale Skalierung acht Kategorien, u. a. Stabilität, Lenkungsrücklauf und Direktheit. Die Autoren führten jedoch selbst als Kritik an, dass diese Anzahl an Aspekten im Widerspruch zur begrenzten Aufnahmekapazität des Menschen steht und vermutlich zu einer bestimmten Zeit nur Teilaspekte des Lenkgefühls wahrgenommen werden können. Insgesamt zeichneten sich sechs häufig genannte Kategorien ab (Abbildung 4). Diese werden nun genauer beschrieben:

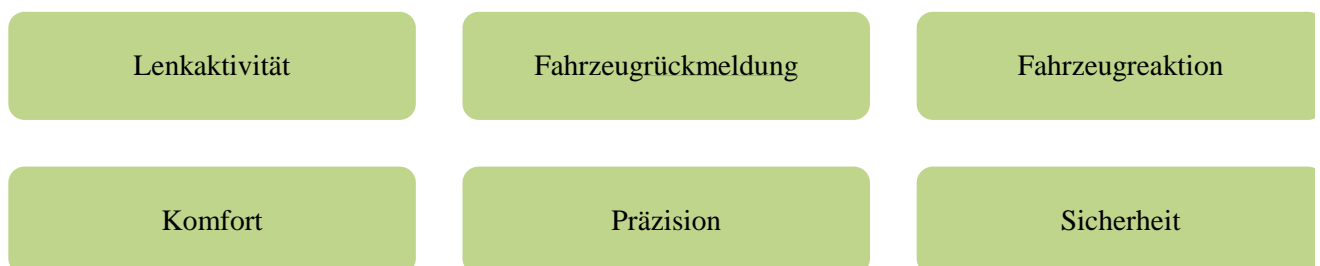


Abbildung 4. Komponenten des Lenkgefühls.

- Die **Lenkaktivität** bzw. der Lenkaufwand beschreibt die Lenkkräfte (Braess, 2001; Mitschke & Wallentowitz, 2014). Diese werden als Lenkmomente gemessen. Teilweise wird auch der benötigte Lenkwinkel dazu in Beziehung gebracht oder zu diesem Aspekt dazugezählt (Wolf, 2009). Unter dem Lenkmoment versteht sich in dieser Arbeit die Höhe des vom Fahrer am Lenkrad aufzuwendenden Moments (Pischinger & Seiffert, 2016). In diesem Zusammenhang wird manchmal auch vom Handmoment gesprochen (Harrer & Pfeffer, 2013; Zschocke, 2009). Mit dem Begriff Lenkwinkel wiederum ist der am Lenkrad gemessene Winkel gemeint (Jürgensohn & Timpe, 2001; Zschocke, 2009).
- Als **Fahrzeugrückmeldung** werden Nutzinformationen und Störinformationen bezeichnet, welche als kleine Lenkradmoment- und Lenkradwinkeländerungen über das Lenkrad an den Fahrer weitergegeben werden (Harrer & Pfeffer, 2013). Nutzinformationen melden dem Fahrer beispielsweise den Fahrzustand zurück und tragen so zu einer angenehmen Regelung bei. Störinformationen, wie durch Fahrbahnunebenheiten hervorgerufene Schläge am Lenkrad, behindern die Fahraufgabe und sollen möglichst reduziert werden.
- Die **Fahrzeugreaktion** folgt auf die Lenkeingabe des Fahrers (Wolf, 2009) und beschreibt den Aufbau der Fahrzeugbewegungen (Harrer & Pfeffer, 2013). Der Zeitraum direkt nach dem Anlenken ist besonders relevant und wird als Ansprechverhalten bezeichnet.
- **Komfort** wird bei Rothhämel (2013) vor allem durch die Abwesenheit von Störungen definiert. Hierunter fallen ruckartige Anregungen oder Vibrationen. Nach Anand (2014) tritt ein hohes Maß an Komfort dann auf, wenn keine besonderen kognitiven Ressourcen nötig sind, um das Lenkmanöver durchzuführen. Dies wird maßgeblich dadurch bestimmt, inwiefern die Erwartungen eines Fahrers aufgrund seiner individuell vorhandenen Fahrerfahrungen erfüllt werden.
- Die **Präzision** oder Zielgenauigkeit wird gemäß Heißing, Ersoy und Gies (2013) dadurch bestimmt, wie störungsfrei ein Fahrzeug dem eingeschlagenen Kurs folgt und ob ein geringer Nachlenkbedarf besteht. Maßgeblich ist hierbei die Passung der Fahrzeugreaktion zum eingestellten Lenkradwinkel (Harrer & Pfeffer, 2013).
- Die empfundene **Sicherheit** spielt nach Mandhata et al. (2012) sowohl bei niedrigen als auch bei hohen Geschwindigkeiten eine Rolle und hängt von der Sensitivität des Lenksystems ab. Nach Ackert (2008) wird das Sicherheitsempfinden maßgeblich durch die haptische Rückmeldung am Lenkrad, beispielsweise über den Fahrbahnzustand, bestimmt.

Aus dieser Vorstellung wird deutlich, dass die Faktoren des Lenkgefühls stark zusammenhängen. Beispielsweise zeigte Anand (2014) mithilfe einer Clusteranalyse, dass Fahrer nicht zwischen den Aspekten Komfort und Kontrollierbarkeit differenzieren und diese gemeinsam einen Faktor bilden. Darüber hinaus ist die Bedeutung der Faktoren für das Lenkgefühl abhängig von der jeweiligen Fahrsituation. So nimmt z. B. die Wichtigkeit der (wahrgenommenen) Stabilität mit steigender Fahrgeschwindigkeit zu und die des Feedbacks ab (Schimmel, Wijts, Jablonowski & Agostini, 2014).

2.3. Stand der Forschung zur Bewertung des Lenkgefühls

Die vorgestellten Komponenten des Lenkgefühls stellen die Basis der subjektiven Bewertung dar. Bevor sie jedoch zur Auslegung und Optimierung von Lenksystemen genutzt werden können, müssen sie entsprechend aufgearbeitet werden. Ein wesentliches Ziel ist es also, die Zusammenhänge zwischen den subjektiven Bewertungen und den objektiv messbaren Parametern aufzudecken. Hierzu werden häufig Fahrversuche durchgeführt, in denen verschiedene Lenkungsauslegungen verglichen werden.

Die wesentlichen Schritte zur Objektivierung sind (Badiru, 2014):

- Auswahl von Fahrzeugen oder Auslegungsvarianten, welche eine breite Spanne an Abstimmungen abdecken
- Objektive Tests zur Messung der objektiven Parameter
- Evaluation der relevanten Aspekte durch Beurteiler anhand von Fragebögen oder Checklisten
- Statistische Analysen

Für die objektiven Tests liegen eine Reihe standardisierter, etablierter Verfahren vor (Gil Gómez et al., 2015; Harrer, Stickel & Pfeffer, 2005). Dies ist jedoch für die subjektive Bewertung nicht der Fall. So unterscheiden sich die Vorgehensweisen beispielsweise in Bezug auf Fahrmanöver, die Nutzung von Referenzfahrzeugen oder die Bewertungsskala. Die Bedingungen, unter denen die Evaluation erfolgt, sind jedoch elementar für die Qualität der subjektiven Bewertungen. Besonders bezüglich der verwendeten Fragebögen ergeben sich einige relevante Fragestellungen, welche nun aus psychologischer Sicht erarbeitet werden (Abbildung 5).

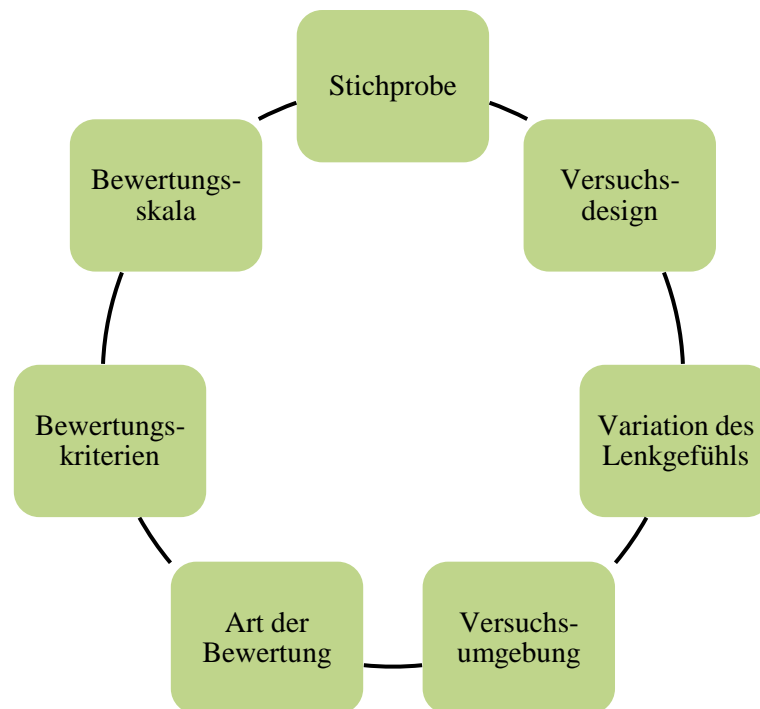


Abbildung 5. Wesentliche Aspekte bei der Konzeption einer Studie zur subjektiven Bewertung des Lenkgefühls.

Zum einen wird die Auswahl der Stichprobe mit besonderem Fokus auf die Expertise der Probanden diskutiert. Weiterhin werden Aspekte des Versuchsaufbaus thematisiert. Hierunter fallen das Versuchsdesign sowie die verschiedenen Möglichkeiten der Variation des Lenkgefühls. Außerdem werden mögliche Versuchsumgebungen und Arten der Bewertung vorgestellt. Schlussendlich werden die in Fragebögen zum Lenkgefühl verwendeten Bewertungskriterien und Bewertungsskalen aufgearbeitet.

2.3.1. Stichprobe

Nachfolgend werden die wesentlichen Aspekte bei der Auswahl einer geeigneten Stichprobe vorgestellt.

Zuerst wird die Expertise bzw. Qualifikation der Probanden diskutiert. In der aktuellen Forschung wird zwischen Experten und Normalfahrern unterschieden (Tabelle 1).

Tabelle 1. Experten und Normalfahrer als Beurteiler des Lenkgefühls.

Experten	Normalfahrer
<ul style="list-style-type: none"> • Beruflich in subjektive Fahrwerksabstimmungen und -bewertungen involviert (Zschocke, 2009) • Vertieftes Verständnis der technischen Zusammenhänge • Geringe Fahrleistung, aber Erfahrung mit verschiedenen Fahrzeugen • Anspruch an Lenkung: Optimierung 	<ul style="list-style-type: none"> • Keine besonderen Fertigkeiten im querdynamischen Grenzbereich (Ullmann, 2006) • Begrenztes Verständnis der technischen Hintergründe • Hohe Fahrleistung, aber fährt häufig nur ein Fahrzeug • Anspruch an Lenkung: Zuverlässigkeit

Die Experten verfügen über Kenntnisse hinsichtlich der Querdynamik sowie Erfahrungen im Bereich der Kraftschlussgrenze. Sie haben aufgrund ihrer beruflichen Tätigkeit ein vertieftes Verständnis zu den technischen Zusammenhängen, die das Lenkgefühl beeinflussen. Außerdem fahren sie im Rahmen des Versuchsbetriebs häufig verschiedene Fahrzeugtypen und Hersteller, weshalb ihnen die Spanne der am Markt verfügbaren Lenksysteme und Lenkungsabstimmungen bekannt ist. Ferner ist ihr Anspruch an die Lenkung, dass sie durch ihre Arbeit immer weiter verbessert werden soll.

Im Gegensatz dazu stehen Normalfahrer, die nach Riedel und Arbinger (1997) weder beruflich noch privat einer ausgiebigen Beschäftigung mit Automobilen nachgehen. Da somit alle Berufskraftfahrer aus dieser Gruppe ausgeschlossen wären, ist diese Definition nicht auf den Lkw anwendbar. Deshalb wird für die vorliegende Arbeit die Definition von Ullmann (2006, S. 3) übernommen, wonach Normalfahrer nicht über „beruflich oder privat begründete Fähigkeiten oder Fertigkeiten bezüglich des Fahrens im querdynamischen Grenzbereich verfügen“. Ihr Verständnis der technischen Hintergründe zur Lenkung ist im Wesentlichen auf die Erhaltung bzw. Wiederherstellung der Fahrbarkeit ausgerichtet. Ferner steuern Normalfahrer aufgrund ihrer primären Tätigkeit als Transporteur meist lediglich ein Fahrzeug bzw. Fahrzeuge eines Herstellers. Somit ist ihnen die Bandbreite der technischen Möglichkeiten nicht bekannt. Im Gegensatz zu Experten weisen sie jedoch deutlich höhere Fahrleistungen auf und bewegen den Lkw regelmäßig in alltäglichen Fahrsituationen. Ihr Anspruch an die Lenkung ist, dass diese zuverlässig funktionieren muss.

In der aktuellen Forschung wird das Heranziehen von Normalfahrern als Beurteiler kontrovers diskutiert. Nach Wolf (2009) sind Normalfahrer, im Gegensatz zu Experten, nicht in der Lage, die wahrgenommenen technischen Unterschiede auf technische Zusammenhänge zurückzuführen. Dies könnte im Wesentlichen in ihrem geringen Verständnis der zu bewertenden Kriterien sowie der fehlenden Übung in der Wahrnehmung der feinen Unterschiede begründet sein (Nybacka, He, Su, Drugge & Bakker, 2014). Außerdem sind Normalfahrer im Vergleich weniger kritisch als Experten bei der Bewertung des Lenkgefühls (Pfeffer & Scholz, 2010). Demgegenüber stehen einige Vorteile bei der Bewertung durch Experten. Ihre Einschätzungen weisen insgesamt

weniger Streuung auf und sind auch bei wiederholter Bewertung (Koch, 2010) oder nahe der Kraftschlussgrenze zuverlässig (Buschardt, 2003). Wegen dieser hohen Qualität der Bewertungen ist außerdem eine niedrigere Stichprobengröße als bei Normalfahrern notwendig. Schlussendlich können Experten insgesamt detailliertere Bewertungen abgeben und einzelne Kriterien besser voneinander differenzieren (Gutjahr, 2013). Diese Argumente führen dazu, dass in einigen Studien lediglich eine sehr kleine Anzahl an Normalfahrern als „Vergleichskollektiv“ mit erhoben wird (Zschocke, 2009, S. 86).

Die Problematiken, die bei Normalfahrern häufig angeführt werden, sind jedoch auch bei Experten zu finden. Chen und Crolla (1998) zeigten, dass die Bewertungen von Experten hinsichtlich der erhobenen Handling-Eigenschaften individuell geprägt waren, was das Aufdecken von Zusammenhängen mit objektiven Parametern erschwert. Dies war mitunter so stark ausgeprägt, dass die Experten keinen mehrheitlichen Konsens erzielten, ob eine experimentelle Bedingung besser war als die Referenz oder nicht. Als mögliche Erklärungen wurden individuelle Vorlieben oder auch individuelle Interpretationen der Bewertungskriterien genannt. Ähnliche Problematiken traten auch bei Gil Gómez et al. (2015) auf. Es liegen auch Belege vor, wonach sogar die Experten innerhalb eines Fahrzeugherstellers entgegen anderer Behauptungen eben kein einheitliches Verständnis bzw. Vokabular zur Beschreibung des Lenkgefühls besitzen (Schimmel et al., 2014). Dies wird damit begründet, dass meist nur eine gesamtheitliche Optimierung des Fahrverhaltens erzielt werden muss aber keine exakte Beschreibung von Detailproblemen. Die Ergebnisse legen den Schluss nahe, dass Experten den Erwartungen, zuverlässig und untereinander stimmig zu bewerten, nicht immer gerecht werden.

Insgesamt ist aufgrund der vorgestellten Studien keine pauschale Aussage zur Eignung von Normalfahrern oder Experten möglich. Vielmehr bringen sowohl Experten als auch Normalfahrer Vor- und Nachteile bzgl. der subjektiven Bewertung des Lenkgefühls mit sich (Koch, 2010). So ist entsprechend der zu untersuchenden Fragestellung zu unterscheiden, welche Qualifikation des Probandenkollektivs optimal ist.

Es gibt jedoch noch weitere Aspekte, welche bei der Zusammenstellung der Stichprobe beachtet werden müssen. Nur mit einer zur Fragestellung passenden Stichprobe können Ergebnisse erzielt werden, welche auf die Realität übertragbar sind. Bei Normalfahrer-Studien werden zur Einordnung des Probandenkollektivs demographische Daten zu Geschlecht, Alter, Fahrleistung und Fahrerfahrung der Teilnehmer berichtet (Anand, 2014; Barthenheier, 2004; Riedel & Arbinger, 1997; Wolf, 2009; Zschocke, 2009). Bei Decker (2009) jedoch finden sich beispielsweise keine Erläuterungen zu Hintergrund und demographischen Daten der Probanden. Andererseits nahmen bei Rothhämel (2013) unter anderem Probanden ohne entsprechenden Führerschein teil. Da diese im Fahren eines Lkw nicht erprobt waren, ist auch hier die Übertragbarkeit der Ergebnisse anzuzweifeln. In diesen Fällen ist also die Eignung der Stichprobe unsicher.

Neben dem entsprechenden Führerschein kann z. B. eine gewisse jährliche Fahrleistung gefordert werden (Dettki, 2005). Ackert (2008) setzte eine Gesamtfahrleistung von mindestens 20.000 km voraus. Ebendiese jährliche Kilometeranzahl gibt Ellmann (2003) als Grenze für Vielfahrer im Pkw an. Fastenmeier (2005) hingegen setzt für erfahrene Fahrer mehr als 100.000 km Fahrpraxis und sechs Jahre Führerscheinbesitz voraus. Bei diesen Grenzwerten ist jedoch zu bedenken, dass sie aus dem Pkw-Bereich stammen und nicht pauschal auf den Lkw

übertragen werden können. Zum einen ist das Handling eines Lkw aufgrund der größeren Maße und Masse schwieriger als das eines Pkw und bedarf mehr Erfahrung. Da Lkw-Fahrer zu beruflichen Zwecken fahren, sind bei ihnen aber ohnehin höhere Jahresfahrleistungen zu finden. Nichtsdestotrotz sind hier auch Unterschiede zwischen den Fahrergruppen zu finden. Während Baustellenfahrer an einem Tag mitunter weniger als 100 km zurücklegen, fährt ein Fernverkehrsfahrer bis zu 800 km. Somit ist ein pauschaler Grenzwert schwer festzulegen.

Genauso bedeutend ist die Festlegung der nötigen Stichprobengröße, da diese ebenfalls entscheidend für die Belastbarkeit der Ergebnisse ist. Bei Bubb (2003) ist zu finden, dass für den Vergleich verschiedener Varianten 30 bis 50 Versuchspersonen nötig sind. Diese Spanne kann jedoch nur als Anhaltspunkt dienen, da die Power eines Tests neben der tatsächlichen Größe des Effekts und dem gewählten Signifikanzniveau von der Stichprobengröße abhängt (Field, 2013). Somit besteht die Möglichkeit, dass selbst große Unterschiede bei kleinen Stichproben zu nicht-signifikanten Ergebnissen führen können. In bisherigen Studien wurde diesen Aspekten wenig Aufmerksamkeit gezollt, sodass teilweise sehr kleine Stichprobengrößen vorlagen. So stützt Zschocke (2009) seine Analysen auf 17 Experten und fünf Normalfahrer. Andere Studien zogen zwölf Experten heran (Nybacka, He, Gil Gómez, Bakker & Drugge, 2014; Nybacka et al., 2014). Wie geschildert ist die Power bei derart kleinen Stichprobengrößen äußerst gering, was zu nicht-signifikanten Ergebnissen führen könnte. Ferner ist auch die Eignung der verwendeten Verfahren, u. a. Varianzanalysen oder Regressionen, zu diskutieren.

Zusammenfassend ist bei der Bildung der Stichprobe die Qualifikation der Probanden, ihre Eignung für die vorliegende Fragestellung sowie die Stichprobengröße zu beachten.

2.3.2. Versuchsaufbau

Versuchsdesign

Bezüglich des Versuchsdesigns können im Wesentlichen between-subject und within-subject Designs unterschieden werden (Renner, Heydasch & Ströhlein, 2012). Bei between-subject Designs werden die verschiedenen Versuchsbedingungen auf die Probanden aufgeteilt (Field, 2013). Das bedeutet, dass später Beurteilungen verschiedener Gruppen miteinander verglichen werden. Bei within-subjects Designs absolviert jeder Proband jede Versuchsbedingung (Bortz & Döring, 2006). So werden später die Beurteilungen eines Probanden untereinander verglichen. Bis auf wenige Ausnahmen, z. B. bei Barthenheier (2004), sind in der Forschung zum Lenkgefühl lediglich within-subject Designs zu finden. Eine solche Messwiederholung ist bei der subjektiven Bewertung anzuraten. Nur so kann die Variation zwischen den Bewertungen nicht durch unterschiedliche Stichproben, sondern nur durch die Variation selbst hervorgerufen worden sein.

Variation des Lenkgefühls

Um das Lenkgefühl zu analysieren, ist es angezeigt, verschiedene Ausprägungen miteinander zu vergleichen. Diese Variation des Lenkgefühls kann durch drei Vorgehensweisen realisiert werden (Tabelle 2).

Erstens kann die Variation durch die Gegenüberstellung der Lenksysteme verschiedener Fahrzeuge realisiert werden. Dieses Vorgehen bietet den Vorteil, dass Serienfahrzeuge genutzt werden können und somit keine zeitintensive Entwicklungsarbeit geleistet werden muss. Jedoch unterscheiden sich die Fahrzeuge in sämtlichen

Eigenschaften und nicht nur in denen, die untersucht werden sollen (Decker, 2009). So können unterschiedliche Sitze mit verändertem Seitenhalt oder unterschiedliche Kranzdurchmesser der Lenkräder zu einer Beeinflussung der Ergebnisse führen. Außerdem erleben die Probanden durch den Wechsel zwischen zwei Fahrzeugen eine Vielzahl an Eindrücken, welche die Erinnerung an das gerade gefahrene Fahrzeug und das erlebte Lenkgefühl abschwächen (Rothhämel, IJkema & Drugge, 2011a). Schlussendlich können auch die optischen Unterschiede der Fahrzeuge die Wahrnehmung und Bewertung der Probanden beeinflussen (Badiru, 2014).

Zweitens kann die Variation durch Veränderungen der Eigenschaften eines Fahrzeugs hervorgerufen werden. Beispielsweise nahmen Riedel und Arbinger (1997) Modifikationen an der Hardware des Fahrzeugs vor, u. a. am Luftdruck der Reifen oder der Dachlast. Dieses Vorgehen hat den Vorteil, dass keine Veränderung des Gesamtfahrzeugs stattfindet und so der Einfluss weiterer Faktoren auf die Bewertung reduziert wird. Allerdings ist die Anzahl an Variablen, die relativ leicht zwischen den Versuchsfahrten geändert werden kann, begrenzt. Schließlich ist bei einer Erprobung im öffentlichen Straßenverkehr auch eine Betriebserlaubnis des Fahrzeugs nötig.

Tabelle 2. Die Vor- und Nachteile der Möglichkeiten zur Variation des Lenkgefühls unter Angabe der Folgen für die Gestaltung des Wechsels.

	Verschiedene Fahrzeuge	Ein Fahrzeug	
		Hardware	Software
Vorteile	<ul style="list-style-type: none"> ➤ Keine zeitintensive Entwicklungsarbeit ➤ Lenkgefühl wird als Teil des Gesamtfahrzeugs erprobt 	<ul style="list-style-type: none"> ➤ Gesamtfahrzeugverhalten bleibt konstant ➤ Veränderungen vergleichsweise einfach realisierbar 	<ul style="list-style-type: none"> ➤ Gesamtfahrzeugverhalten bleibt konstant ➤ Zahlreiche Veränderungsmöglichkeiten
Nachteile	<ul style="list-style-type: none"> ➤ Unterschiedliches Gesamtfahrzeugverhalten ➤ Unterschiedliche Optik, Ausstattung, etc. 	<ul style="list-style-type: none"> ➤ Anzahl veränderbarer Variablen begrenzt 	<ul style="list-style-type: none"> ➤ Eingriff ins Lenksystem ➤ Ggf. zusätzliche Freigaben nötig
Wechsel	<ul style="list-style-type: none"> ⊖ Proband muss umsteigen ⊖ Umstieg stellt Unterbrechung dar 	<ul style="list-style-type: none"> ⊖ Proband muss zum Umbau aussteigen ⊖ Ausstieg stellt Unterbrechung dar 	<ul style="list-style-type: none"> ⊕ Ohne Unterbrechung während Fahrt ⊕ Wechsel kann auch für den Proband unsichtbar erfolgen ⊕ Proband kann selbst umschalten

Viel häufiger jedoch erfolgt die Variation der Software-Komponenten innerhalb eines Fahrzeugs, indem verschiedene Parametrierungen des Lenksystems implementiert werden. Das bedeutet, dass die der Lenkunterstützung zugrunde liegenden Kennlinien systematisch variiert werden. So können die Kennwerte der Lenkungsauslegung gezielt verändert werden, während das Gesamtfahrzeugverhalten und andere Parameter

konstant bleiben. Diese Variation ist während der Fahrt innerhalb von wenigen Sekunden möglich (Decker, 2009). So ist der Wechsel zwischen Lenkungsvarianten nicht mit dem Umsteigen in ein anderes Fahrzeug oder dem Umrüsten von Hardware-Komponenten verbunden, wodurch der subjektive Vergleich erleichtert wird. Dieses Vorgehen verlangt jedoch einen Eingriff in das Lenksystem, sodass häufig nur noch Testfahrten auf Prüfgeländen möglich sind.

In Abhängigkeit der gewählten Variation und der zu beantwortenden Fragestellung sind verschiedene Vorgehensweisen möglich, wie zwischen den einzelnen Varianten gewechselt wird.

Bei der Verwendung verschiedener Fahrzeuge ist es zwingend nötig, dass der Proband umsteigt und so die Veränderung unmittelbar wahrnimmt. Bei der Verwendung nur eines Fahrzeugs und der Variation von Hardware-Komponenten ist es ebenso unumgänglich, dass der Proband den Zeitpunkt der Veränderung wahrnimmt. Er kann lediglich über die Art der Veränderung im Unklaren gelassen werden, sofern diese nicht von außen ersichtlich ist (Riedel & Arbinger, 1997). Auch bei der Variation von Software-Komponenten wird der Fahrer meist über den Wechsel zwischen den Varianten informiert (Harrer, 2007; Rothhämel, 2013). Der Wechsel findet dann zu einem festgelegten Zeitpunkt während oder nach dem Durchfahren eines Manövers statt. Mit diesem Vorgehen können sich die Probanden auf die Umstellung einstellen und zum entsprechenden Zeitpunkt gezielt auf Veränderungen achten.

Bei der Variation von Software-Komponenten ist es jedoch auch möglich, den Probanden über das Umschalten im Unklaren zu lassen. So kann beispielsweise untersucht werden, ob Probanden Unterschiede wahrnehmen, wenn eigentlich keine Veränderung seitens des Versuchsleiters vorgenommen wurde (Barthenheier, 2004). Es ist außerdem möglich, den Testpersonen selbst die Möglichkeit zum Umschalten zu geben, beispielsweise über Knöpfe am Lenkrad (Anand, Terken & Hogema, 2011). Dieses Vorgehen hat den Vorteil, dass die Probanden die Dauer der Evaluation der einzelnen Varianten selbst bestimmen können und diese nicht durch den Wechsellpunkt vorgegeben ist. Ein etwas abgeschwächtes Vorgehen ist bei Decker (2009) zu finden, bei dem Probanden beliebig oft zwischen den zu vergleichenden Varianten wechseln konnten.

Versuchsumgebung

Wie aus der Usability-Forschung bekannt, können Menschen ihr Wissen zu Produkten und deren Handhabung besonders gut abrufen, wenn sie sich in den entsprechenden Nutzungssituationen befinden (Richter & Flückiger, 2013). Somit kann in einem ersten Schritt festgehalten werden, dass die Lenkung im Fahrzeug und während der Fahrt erprobt werden muss. Im nächsten Schritt muss dann die für die jeweilige Fragestellung geeignete Versuchsumgebung identifiziert werden. Es können im Wesentlichen drei verschiedene Testumgebungen für die Lenkungsbewertung unterschieden werden (Tabelle 3):

Tabelle 3. Die Vor- und Nachteile der zur Bewertung des Lenkgefühls geeigneten Versuchsumgebungen.

	Öffentlicher Straßenverkehr	Testgelände	Fahrsimulator
Vorteile	<ul style="list-style-type: none"> • Realistische Umgebung, natürliches Fahrverhalten • Höchste Validität, da Bewertungsgegenstand unter realen Bedingungen getestet werden kann (Breuer, 2000) 	<ul style="list-style-type: none"> • Realitätsnahe Umgebung • Reproduzierbarkeit der zu bewertenden Szenarien • Versuchsablauf stark standardisierbar • Systematische Variation von Faktoren leicht realisierbar 	<ul style="list-style-type: none"> • Hohe Reproduzierbarkeit der Versuche und Unabhängigkeit von äußeren Faktoren, z. B. Witterung (Kraft, 2011) • Viele Freiheitsgrade bei der Gestaltung der Versuchsvarianten (Dang et al., 2015) • Systematische Variation von Faktoren leicht realisierbar (Zöller, 2015)
Nachteile	<ul style="list-style-type: none"> • Systematische Variation von Faktoren sehr aufwendig, da der Versuchsablauf an örtliche Gegebenheiten angepasst werden muss (Breuer, 2000) • Reales Verkehrsgeschehen bringt Einflüsse durch andere Verkehrsteilnehmer mit sich (Bruder, Abendroth & Landau, 2007) 	<ul style="list-style-type: none"> • Begrenzte Fläche, begrenzter Geschwindigkeitsbereich 	<ul style="list-style-type: none"> • Realität wird aufgrund fehlender Tiefenhinweise und eingeschränktem haptischen Feedback nicht vollständig abgebildet (Knappe, 2009) • Veränderungen des Fahrverhaltens durch eingeschränktes Gefährdungsbewusstsein (Breuer, 2000) • Simulatorkrankheit (Caird & Horrey, 2011)

Zum einen kann eine Erprobung auf einer festgelegten Strecke im öffentlichen Straßenverkehr stattfinden. Beispiele hierfür sind Arbeiten von Riedel und Arbinger (1997) sowie Wolf (2009). Diese Versuchsumgebung ist am realistischsten, da hier unter realen Bedingungen getestet werden kann. So ist keine Gewöhnung der Probanden an unbekannte Versuchsumgebungen nötig. Allerdings können die Versuchsbedingungen nicht stabil gehalten werden, da beispielsweise die Verkehrslage eine Anpassung der Fahrgeschwindigkeit verlangt. Außerdem sind die Versuchsstrecke sowie der Wechsellpunkt an die Straßenführung und Platzverhältnisse anzupassen. Diese Versuchsumgebung ist somit vor allem für einen ersten Eindruck sowie eine globale Bewertung des Lenkgefühls geeignet.

Es finden sich jedoch auch Simulatorstudien zum Lenkgefühl (Anand, 2014; Rothhämel, 2013). Hier ist zum einen eine hohe Reproduzierbarkeit gegeben, da die Umweltfaktoren komplett kontrolliert werden können. Außerdem ist hier im Gegensatz zum öffentlichen Straßenverkehr der Wechsel zwischen verschiedensten Varianten per Knopfdruck möglich. Probleme ergeben sich jedoch zum einen in der begrenzten Realitätstreue der Simulation. Zum anderen ist von einem veränderten Fahrverhalten der Probanden, u. a. aufgrund der veränderten Wahrnehmung von Gefahren auszugehen. So zeigte Zöller (2015), dass sich die Standardabweichungen des

Lenkradwinkels zwischen der Realfahrt und dem Fahrsimulator unterschieden, wobei in den meisten Fällen im Fahrsimulator höhere Werte auftraten.

Schlussendlich können Teststrecken zur Durchführung von Fahrversuchen genutzt werden (Gil Gómez et al., 2015; Nybacka et al., 2014). Auf diesen können verschiedene Parametrierungen des Lenksystems gezielt erlebbar gemacht und so beurteilt werden. Zschocke (2009) und Harrer (2007) nutzten für die Lenkungsbewertung durch Experten ISO-Manöver wie beispielsweise den doppelten Spurwechsel (ISO 3888-1:1999) oder den Weave-Test (ISO 13674-1:2010). Diese ermöglichen eine standardisierte Erprobung, sind jedoch aufgrund ihrer hohen Querschleunigungen und Ansprüche an das fahrerische Können für Normalfahrer teilweise nicht geeignet. Bei Evaluationen durch Normalfahrer bieten sich mehrere Möglichkeiten an. Einerseits können die ISO-Manöver abgeschwächt werden, um sie für Normalfahrer beherrschbar zu machen. Andererseits kann die Evaluation im Rahmen eines vorgegebenen Kurses auf vorhandenen Teststrecken des Prüfgeländes wie z. B. einem Landstraßenkurs erfolgen. Außerdem ist auf dem Prüfgelände eine freie Exploration möglich (Chen, 1997; Nybacka et al., 2014). Auch Gil Gómez et al. (2015) gaben Experten die Möglichkeit, Fahrzeuge völlig frei zu erproben. Bei dieser Herangehensweise zeigte sich jedoch, dass Fahrzeuge allein aufgrund unterschiedlicher Fahrweisen unterschiedlich bewertet werden. Anders als im Straßenverkehr ist auf der Teststrecke der Wechsel zwischen verschiedenen Fahrzeugen oder Varianten zwischen den einzelnen Manövern einfacher möglich. Jedoch sind die Möglichkeiten auch auf der Teststrecke durch die zur Verfügung stehende Fläche begrenzt, sodass die nötigen Fahrgeschwindigkeiten aufgrund des begrenzten Platzes möglicherweise nicht umsetzbar sind. Somit muss im Einzelfall geprüft werden, inwiefern sich eine Teststrecke zur Durchführung der geplanten Fahrmanöver eignet.

Insgesamt muss entsprechend der Fragestellung entschieden werden, welche Versuchsumgebung am geeignetsten ist (Bruder et al., 2007). Ackert (2008) resümiert dazu, dass in den vorliegenden Berichten keine Aussagen dazu zu finden sind, welche Versuchsumgebung bei welcher Fragestellung am geeignetsten ist.

Art der Bewertung

Unter der Art der Bewertung wird in dieser Arbeit die Möglichkeit zum Vergleich zwischen den angebotenen Varianten bzw. Fahrzeugen verstanden. Hier ist zwischen absoluten und relativen Bewertungen zu unterscheiden (Bruder, Fuchs, Abendroth & Schramm, 2009).

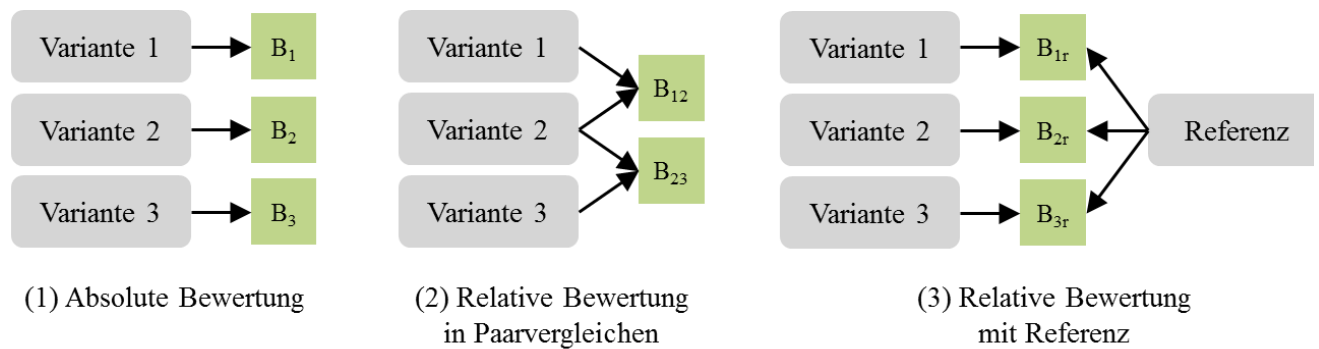


Abbildung 6. Absolute und relative Bewertung (B) am Beispiel drei bewerteter Varianten.

Unter absoluten Bewertungen (siehe (1) in Abbildung 6) versteht man, dass die Bewertungen unabhängig von anderen Fahrzeugen, Lenksystemen o. ä. abgegeben werden. Sie haben deshalb den Vorteil, dass sie verallgemeinerbar und auch isoliert betrachtet aussagekräftig sind.

Bei relativen Bewertungen wird die Bewertung in Bezug auf andere Fahrzeuge oder Varianten abgegeben. Zum einen kann die Gegenüberstellung in Paarvergleichen erfolgen (siehe (2) in Abbildung 6). Das bedeutet, dass zwei bestimmte Varianten gegeneinander verglichen werden (Barthenheier, 2006). Häufig wird als Bezugspunkt auch eine sogenannte Referenz benutzt (siehe (3) in Abbildung 6). Diese wird oft vor den eigentlichen Fahrzeugen oder Varianten bewertet (Harrer, 2007) und bzw. oder immer wieder während des Versuchs wiederholt (Breuer, 2000). Auf diese Weise soll verhindert werden, dass insbesondere Normalfahrer einen individuellen Bezugspunkt heranziehen. Dies könnte beispielsweise ihr üblicherweise gefahrenes Fahrzeug sein.

In einigen Studien zeigte sich, dass absolute Bewertungen von Normalfahrern nicht reproduzierbar waren (Barthenheier & Winner, 2003; Rothhämel, 2013). Relative Bewertungen hingegen führten besonders bei Normalfahrern zu kleineren Streuungen (Buschardt, 2003). Durch die Nutzung einer Referenz kann verhindert werden, dass sich der individuelle Anker einer Person während der Versuchsdauer verschiebt (Breuer, 2000). Somit sind für die Bewertung des Lenkgefühls durch Normalfahrer relative Beurteilungen heranzuziehen (Gies & Marusic, 2000).

Zusammenspiel aus Variation des Lenkgefühls, Versuchsumgebung und Art der Bewertung

Der Versuchsaufbau und -ablauf ist insbesondere im Hinblick auf einen möglichst unmittelbaren Vergleich der Varianten bzw. Fahrzeuge auszulegen. Dies wird im Wesentlichen durch die Variation des Lenkgefühls und die gewählte Versuchsumgebung bestimmt. Die Veränderung des Lenkgefühls kann durch die Nutzung verschiedener Fahrzeuge oder durch Hardware- oder Software-Änderungen an einem Fahrzeug hervorgerufen werden. Als Versuchsumgebung kann der öffentliche Straßenverkehr, das Testgelände oder der Fahrsimulator gewählt werden. Aus der gewählten Kombination ergibt sich dann, ob eine relative Bewertung möglich ist oder eine absolute Bewertung genutzt werden muss.

Die mögliche bzw. empfohlene Art der Bewertung ist zusammen mit den Vor- und Nachteilen der jeweiligen Kombination in Tabelle 4 dargestellt. Bei relativen Bewertungen ist es wichtig, dass der zeitliche Abstand

zwischen den zu vergleichenden Varianten möglichst kurz ist. Deshalb ist diese Form des Vergleichs nicht bei allen Versuchskonzepten möglich. Die relative Bewertung und unmittelbare Vergleichbarkeit ist besonders bei der Variation des Lenkgefühls über Hardware-Änderung und die Nutzung verschiedener Fahrzeuge erschwert, da hier eine Unterbrechung der Bewertung für den Wechsel zwischen Varianten nötig ist. Dies kann bei Versuchen im Straßenverkehr aufgrund nötiger Transferfahrten mit einem hohen zeitlichen Aufwand verbunden sein. Nichtsdestotrotz muss auch bei Versuchen auf der Teststrecke oder im Fahrsimulator Augenmerk darauf gelegt werden, die Vergleichbarkeit durch den Versuchsablauf zu maximieren.

Tabelle 4. Zusammenstellung der Vor- und Nachteile sowie die mögliche Art der Bewertung in Abhängigkeit der Versuchsumgebung und der Variation des Lenkgefühls.

Variation des Lenkgefühls	Versuchsumgebung		
	Öffentlicher Straßenverkehr	Testgelände	Fahrsimulator
Verschiedene Fahrzeuge	<ul style="list-style-type: none"> ⊕ Einfach umsetzbar ⊖ Beeinflussung durch Gesamtfahrzeug und Varianz in Versuchsumgebung ➤ Art der Bewertung: absolut 	<ul style="list-style-type: none"> ⊕ Lenkgefühl als Teil des Gesamtfahrzeugverhaltens standardisiert evaluierbar ⊖ Kein direkter Vergleich möglich ➤ Art der Bewertung: relativ, falls nötig absolut 	<ul style="list-style-type: none"> ⊕ Einfacher Wechsel zwischen simulierten Fahrzeugen ⊖ Lenkrad sowie Fahrerhaus nicht austauschbar ➤ Art der Bewertung: relativ
Verschiedene Varianten	<ul style="list-style-type: none"> ⊕ Realitätsnahe, blinde Erprobung ⊖ Schwer umsetzbar, da Straßenfreigabe nötig ➤ Art der Bewertung: relativ 	<ul style="list-style-type: none"> ⊕ Unmittelbarer Vergleich in standardisierten Fahrmanövern ⊖ Evtl. Ergebnisse schwer übertragbar ➤ Art der Bewertung: relativ 	<ul style="list-style-type: none"> ⊕ Variation leicht umsetzbar ⊖ Ergebnisse von Qualität des Fahrsimulators abhängig ➤ Art der Bewertung: relativ

2.3.3. Fragebogen

Bewertungskriterien

Wie bereits angesprochen, basiert die subjektive Bewertung des Lenkgefühls auf Checklisten und Fragebögen. Diese beinhalten nach Themenbereichen sortierte Bewertungskriterien, welche das Lenkgefühl insgesamt oder nur einen bestimmten Teil erfassen. Zuerst ist anzumerken, dass nur wenige Arbeiten auf den Lkw ausgerichtet sind (Rothhämel, 2013; Rothhämel et al., 2010, Rothhämel et al., 2011a, 2011b). Die auf den Pkw ausgerichteten Fragebögen unterscheiden sich dabei sehr stark hinsichtlich der Quantität und Tiefe der Bewertungskriterien. In Tabelle 5 wird ein Überblick über die Fragebögen ausgewählter Studien gegeben.

Tabelle 5. Überblick zu Bewertungskriterien ausgewählter Fragebögen unter Angabe der Zielgruppe.

Studie	Zielgruppe*	Beispielkriterium	Anzahl Kriterien
Kim und Yoon (2015)	EXP	Zeitverzug**	11
Mandhata et al. (2012)	NF	Ich musste viel Kraft aufwenden, um das Fahrzeug dorthin zu bekommen, wo ich wollte.**	9
Rothhämel et al. (2010)	NF / S	Leicht / schwer**	32
Wolf (2009)	NF	Agilität: Verhält sich das Fahrzeug ausreichend agil oder ist es etwas träge?	23
Zschocke (2009)	NF / EXP	Lenkmomentgradient im Weiterlenkbereich: Lenkkraftanstieg über Querschleunigungsbereich	27
Harrer (2007)	EXP	Lenkwinkelgeschwindigkeit	26
Riedel und Arbinger (1997)	NF	Das Fahrzeug reagierte direkt auf die Lenkung.	7

* EXP: Experten / NF: Normalfahrer / S: Sonstige

** Übersetzung aus dem Englischen

Zum einen sind die Fragebögen dahingehend zu unterscheiden, für welche Beurteiler sie angewendet werden sollen. Fragebögen für Experten zeichnen sich durch die Verwendung von Fachbegriffen und ihre Detailliertheit aus. Ferner ist eine gewisse Erfahrung nötig, um die beschriebenen Eigenschaften wahrnehmen zu können. Diese Aspekte gehen aus den Beispielkriterien der Expertenfragebögen (Harrer, 2007, Kim & Yoon, 2015, 2015; Zschocke, 2009) deutlich hervor. Die Verwendung solcher Kriterien bei Normalfahrern beeinflusst die Qualität der Bewertungen, da die Kriterien nicht vollständig verstanden werden können. Dies spiegelt sich beispielsweise bei Zschocke (2009) in den wenig aussagekräftigen Ergebnissen der Normalfahrer wieder.

Bei Fragebögen für Normalfahrer ist daher darauf zu achten, dass sie ihrer Expertise entsprechen (Gies & Marusic, 2000). Deswegen sind die Kriterien meist allgemeiner gehalten und beinhalten keine oder kaum spezifische Fachbegriffe (Riedel & Arbinger, 1997; Rothhämel et al., 2010; Wolf, 2009). Dies führt zu einer entsprechend reduzierten Auflösung des Lenkgefühls sowie globaleren Bewertungen. Da Experten so nicht ihr gesamtes Potential ausnutzen können und die Detailliertheit der Ergebnisse unnötig reduziert wird, ist die Verwendung eines Normalfahrerfragebogens bei Experten nicht ratsam.

Ein weiterer wichtiger Aspekt zur Beschreibung eines Fragebogens ist die Anzahl an Bewertungskriterien. Diese variiert stark zwischen sieben Kriterien bei einem Fragebogen für Normalfahrer (Riedel & Arbinger, 1997) und 26 in einem Expertenfragebogen (Harrer, 2007). Bei anderen Ansätzen wird den Probanden zu Beginn des Versuchs eine größere Anzahl an Kriterien vorgelegt, wobei später nur die darüberstehenden Skalen bewertet werden (Rothhämel et al., 2011a). Somit besteht ein Spannungsfeld zwischen Versuchsökonomie einerseits und

umfassenden Ergebnissen andererseits. Zwar führt eine größere Anzahl an Kriterien zu einem detaillierteren Bild. Allerdings muss die Anzahl an Kriterien handhabbar bleiben, insbesondere wenn mehrere Varianten verglichen werden und somit der Fragebogen von jedem Probanden mehrfach beantwortet wird. In diesem Zusammenhang ist die Ermüdung der Probanden über die Versuchsdauer hinweg ebenfalls nicht zu vernachlässigen.

Allgemein ist anzumerken, dass nur selten Angaben dazu zu finden sind, wie der genutzte Fragebogen entstanden ist und validiert wurde. In einigen Studien wurde mithilfe von Interviews ein erster Pool an Bewertungskriterien gesammelt (Anand, 2014; Riedel & Arbinger, 1997; Rothhämel et al., 2010). Anschließend wurden die Fragebögen auf Basis von Clusteranalysen oder Fahrversuchen optimiert. Eine weitere, leicht realisierbare Herangehensweise zur Datengewinnung und darauf aufbauenden Optimierung des Fragebogens können auch online-Befragungen sein (Black, 2010). Dennoch ist eine Einschätzung der Qualität eines Fragebogens ohne Angaben zum Entwicklungsprozess nicht möglich.

Der aktuelle Forschungsstand lässt also den Schluss zu, dass es bezüglich der Bewertung des Lenkgefühls am Lkw derzeit an validierten Fragebögen fehlt.

Bewertungsskala

Neben den Bewertungskriterien selbst muss eine für die Fragestellung geeignete Skala gefunden werden. In der psychologischen Forschung wird die Bedeutung der Skala vielfach hervorgehoben, da sie einen maßgeblichen Einfluss auf die psychometrischen Gütekriterien hat (Kallus, 2010).

Allgemein kann zwischen offenen und geschlossenen Fragen unterschieden werden. Bei offenen Fragen ist die Antwort gänzlich frei und wird vom Probanden verbal oder schriftlich gegeben. Ein Beispiel hierfür wäre: „Was fällt Ihnen zum Thema Lenkung ein?“. Der Vorteil offener Fragen ist, dass der Proband ohne Einschränkungen antworten kann. Wesentliche Nachteile sind jedoch die geringe Auswertungsobjektivität bei gleichzeitig hohem Auswertungsaufwand (Moosbrugger & Kelava, 2012). In der Forschung zum Lenkgefühl wurde dieses Skalenformat beispielsweise von Wolf (2009) bei der Evaluation des Lenkgefühls verschiedener Fahrzeuge genutzt. Die freien Antworten der Probanden wurden dann vom Versuchsleiter in eine 3-stufige Skala übertragen.

Insgesamt setzte sich in der Forschung zum Lenkgefühl das geschlossene Fragenformat durch. Dieses zeichnet sich dadurch aus, dass eine gewisse Anzahl an Antwortmöglichkeiten bzw. Stufen vorgegeben ist (Abbildung 7). Die Antwortalternativen variieren dabei hinsichtlich Häufigkeit, Intensität oder Grad der Zustimmung. Die Daten, welche mithilfe dieser sogenannten Likert-Skalen (Likert, 1932) gesammelt werden, lassen auch die Berechnung interferenzstatistischer Verfahren wie *t*-Tests oder Varianzanalysen zu.

(1) 5-stufige Absolut-Skala

	1	2	3	4	5
Die Lenkung ist leichtgängig.					
	gar nicht				völlig

(2) 7-stufige Relativ-Skala

	-3	-2	-1	0	1	+2	+3
Wie leichtgängig ist die Variante im Vergleich zur Referenz?							
	viel weniger						viel mehr

(3) 5-stufige, benannte Absolut-Skala

	1	2	3	4	5
Die Lenkung ist leichtgängig.					
	gar nicht	kaum	eher	ziemlich	völlig

(4) Semantisches Differential

	1	2	3	4	5
Die Lenkung ist...					
	leichtgängig				schwergängig

Abbildung 7. Skalenformate.

Ein wesentliches Merkmal solcher Likert-Skalen ist die Anzahl der angebotenen Stufen, da sie den Detailgrad der Antworten bestimmt (vergleiche (1) und (2) in Abbildung 7). In der Forschung zum Lenkgefühl sind zwischen vier (Riedel & Arbinger, 1997) und zehn (Zschocke, 2009) Antwortmöglichkeiten zu finden (Tabelle 6). Je nach Fragestellung muss außerdem diskutiert werden, inwiefern eine Mittelkategorie sinnvoll und somit eine Skala mit ungerader Anzahl an Skalenpunkten zu wählen ist. Bei einer relativen Bewertung ist die Nutzung einer ungeraden Anzahl an Skalenpunkten sinnvoll (Rothhämel, 2013). Schließlich kann durch die Auswahl der Mittelkategorie abgebildet werden, dass kein Unterschied im Vergleich zur Referenz wahrgenommen wurde (siehe (2) in Abbildung 7). In jedem Fall sollte die Anzahl an Skalenpunkten über alle Bewertungen hinweg konstant sein. Besonders bei Normalfahrern sollte, im Gegensatz zu Harrer (2007), Zschocke (2009) oder Dang et al. (2015), der Wechsel zwischen Skalen mit unterschiedlicher Anzahl an Skalenpunkten vermieden werden. Dies erleichtert dem Probanden, das Bewertungsvorgehen zu verinnerlichen. Ein weiterer Aspekt ist die verbale Benennung von Kategorien (Porst, 2011) (siehe (3) in Abbildung 7). Diese kann für alle Skalenpunkte (Riedel & Arbinger, 1997), einige Skalenpunkte (Zschocke, 2009) oder nur die Pole (Mandhata et al., 2012) vorgenommen werden.

Im Rahmen der subjektiven Bewertung werden auch semantische Differentiale (Osgood, 1952, 1964) (siehe (4) in Abbildung 7) genutzt, beispielsweise bzgl. des Schwingungskomforts (Jürgensohn & Kolrep, 2006). Im Bereich der Bewertung des Lenkgefühls werden sie von Rothhämel et al. (2011a) eingesetzt. Bei diesem Format werden keine direkten Fragen verwendet. Stattdessen stehen sich an den Polen der Skala bestimmte Eigenschaftspaare gegenüber, beispielsweise *leichtgängig* und *schwergängig*. Die Bewertung wird dann über die Position der gewählten Antwortkategorie zwischen den beiden Gegensatzpolen abgegeben. In Anlehnung an dieses Skalenformat ist auch eine Benennung der Pole mit Eigenschaftsworten, die zum bewerteten Kriterium passen, möglich (Harrer, 2007; Kim, 2011; Kim & Yoon, 2015; Zschocke, 2009).

Tabelle 6. Eigenschaften der Skalen ausgewählter Fragebögen unter Angabe der Zielgruppe.

Studie	Zielgruppe*	Art der Bewertung	Mehrstufige Bewertung	Anzahl der Stufen	Benennung
Kim und Yoon (2015)	EXP	Absolut	<ul style="list-style-type: none"> • Niveau • Gefallen 	10	Alle Stufen
Mandhata et al. (2012)	NF	Absolut	-	7	Pole
Rothhämel et al. (2010)	NF / S	Relativ	-	7	Pole und Mitte (Semantisches Differential)
Wolf (2009)	NF	Absolut	-	Frei / 3	-
Zschocke (2009)	NF / EXP	Absolut	<ul style="list-style-type: none"> • Niveau • Gefallen 	<ul style="list-style-type: none"> • 6 • 10 	<ul style="list-style-type: none"> • Niveau: Pole • Gefallen: einige Stufen
Harrer (2007)	EXP	Absolut	<ul style="list-style-type: none"> • Niveau • Abweichung vom Optimum 	<ul style="list-style-type: none"> • 10 • 4 / 7 	<ul style="list-style-type: none"> • Niveau: alle Stufen • Abweichung vom Optimum: Pole und Mitte
Riedel und Arbinger (1997)	NF	Absolut	-	4	Alle Stufen

* EXP: Experten / NF: Normalfahrer / S: Sonstige

Speziell zur Bewertung des Lenkgefühls sind zwei weitere Herangehensweisen verbreitet. Erstens wird die Bewertung des Lenkgefühls in mehreren Studien in zwei Stufen aufgeteilt (Kim, 2011; Kim & Yoon, 2015; Zschocke, 2009). Zuerst wird bewertet, inwiefern eine bestimmte Eigenschaft auf das gerade gefahrene Fahrzeug bzw. die gerade gefahrene Variante zutrifft. Für diese quasi-objektiven Niveau-Bewertungen (Decker, 2009) fungieren die Fahrer als Sensor und sollten deshalb zu vergleichbaren Bewertungen kommen. Im zweiten Schritt schätzen die Probanden ein, wie gut ihnen das wahrgenommene Verhalten bezüglich des vorliegenden Kriteriums gefällt. Diese Gefallen-Bewertungen sind individuell und hängen von den persönlichen Präferenzen ab. Einige Fahrer bevorzugen beispielsweise generell niedrigere Lenkmomente, andere höhere. Ferner sind die Bewertungen

von Niveau und Gefallen abhängig von der Fahrsituation (Gil Gómez et al., 2015). So wird beim Parkieren ein geringes aufzuwendendes Lenkmoment bevorzugt, bei der Autobahnfahrt jedoch ein höheres. Bei Harrer (2007) und Decker (2009) kann abschließend bei auffälligen Bewertungen außerdem die Richtung der Abweichung vom Ideal angegeben werden. Unabhängig davon, welches Vorgehen verfolgt wird, sollte die Anzahl an Bewertungsschritten über die gesamte Evaluation konstant sein. Ein inkonsistenter Bewertungsprozess, wie beispielsweise bei Decker (2009), birgt die Gefahr der Konfusion auf Seiten der Probanden.

Zweitens wird häufig die in der Automobilindustrie geläufige 10-stufige SAE-Skala verwendet oder als Grundlage herangezogen (Harrer, 2007; Heißing & Brandl, 2002; Zschocke, 2009). Der untere Pol der Skala steht dabei für ein Sicherheitsrisiko, der obere Pol für eine hervorragende Bewertung (Harrer & Pfeffer, 2013). Die Bedeutung der Skalenpunkte wird zusätzlich durch einen Farbverlauf nach dem Ampel-Schema unterstrichen. Durch diese Farbkodierung wird jedoch die der Intervallskalierung zugrunde liegende Annahme, dass die Abstände zwischen den einzelnen Skalenpunkten gleich groß sind, verletzt. Schließlich ist anzunehmen, dass der Abstand zwischen zwei Punkten mit unterschiedlichen Farben größer wahrgenommen wird als der zwischen zwei Punkten der gleichen Farbe. Außerdem fanden Gil Gómez et al. (2015), dass die meisten Bewertungen zwischen 6 und 8 liegen und somit nur ein Bruchteil der Skala ausgenutzt wird.

Aus der Summe der vorgestellten Skalen wird ersichtlich, dass sich bislang noch kein Antwortformat durchsetzen konnte. Vielmehr scheint eine individuelle Prüfung des Anwendungsfalls nötig zu sein, um die am besten geeignetste Skala zu identifizieren.

2.4. Erkenntnisse zur Bewertung des Lenkgefühls

Im vorherigen Kapitel wurde ein Überblick gegeben, wie die subjektive Bewertung des Lenkgefühls erfolgen kann. Es wurde herausgearbeitet, dass hierzu Fahrversuche mit einer zur Fragestellung passenden Stichprobe nötig sind. Aus der gewählten Variation des Lenkgefühls und der Versuchsumgebung ergibt sich, welche Art der Bewertung genutzt werden kann. Schlussendlich müssen Anforderungen an Tiefe und Anzahl der Bewertungskriterien sowie die Skala abgeleitet werden.

In Tabelle 7 ist eine Auflistung relevanter Studien zum Lenkgefühl zu finden. Die meisten von ihnen sind auf den Pkw ausgerichtet. Dabei findet eine Aufteilung nach Art der Variation des Lenkgefühls und der Versuchsumgebung statt. Ferner wird die verwendete Stichprobe angegeben. Auf die Studien, die nicht ausschließlich Experten als Beurteiler heranzogen, wird in den folgenden Unterkapiteln näher eingegangen.

Tabelle 7. Studien zum Lenkgefühl in Abhängigkeit der Versuchsumgebung, Variation des Lenkgefühls und Stichprobe*.

Variation des Lenkgefühls	Versuchsumgebung		
	Straßenverkehr	Testgelände	Fahrsimulator
Mehrere Fahrzeuge	<ul style="list-style-type: none"> • Wolf (2009): NF 	<ul style="list-style-type: none"> • Gil Gómez et al. (2015): EXP • Kim und Yoon (2015): EXP • Rothhämel (2013): S • Kim (2011): EXP • Zschocke (2009): NF / EXP • Harrer (2007): EXP 	
Mehrere Varianten	<ul style="list-style-type: none"> • Riedel und Arbinger (1997): NF / EXP 	<ul style="list-style-type: none"> • Anand (2014): NF • Decker (2009): NF • Ackert (2008): NF • Barthenheier (2004): NF 	<ul style="list-style-type: none"> • Dang et al. (2015): NF / EXP • Anand (2014): NF • Mandhata et al. (2012): NF • Rothhämel et al. (2011a): S • Anand et al. (2011): NF • Black (2010): NF

* EXP: Experten / NF: Normalfahrer / S: Sonstige

2.4.1. Studien im Straßenverkehr

Fahrversuche im öffentlichen Straßenverkehr zeichnen sich durch reale Fahrsituationen aus, wodurch Ergebnisse mit einer hohen externen Validität erzielt werden können. Insgesamt liegen nur wenige Studien zur subjektiven Bewertung des Lenkgefühls in dieser Versuchsumgebung vor. Bei den zwei vorgestellten Arbeiten wurde eine Variation des Lenkgefühls durch die Nutzung verschiedener Fahrzeuge oder die Veränderung von Hardware-Parametern erreicht (Riedel & Arbinger, 1997; Wolf, 2009). In beiden Studien werden jedoch die Grenzen dieses Ansatzes sichtbar. Ein Überblick über die verwendeten Methoden ist in Tabelle 8 zu finden.

Bei Wolf (2009) wurde die Wichtigkeit einzelner Aspekte des Lenkgefühls in Bezug auf verschiedene Fahrzeugtypen analysiert. Aus den Bewertungen des Lenkgefühls verschiedener Fahrzeuge gingen hauptsächlich für die Aspekte des Lenkmoments Unterschiede hervor. Hierin wird erneut die zentrale Bedeutung des Lenkmoments für das Lenkgefühl sichtbar. Zu den aufgedeckten Zusammenhängen zwischen den subjektiven Bewertungen der Fahrer und den objektiven Parametern liegen lediglich qualitative Ergebnisse vor, sodass nur grundlegende Auslegungsempfehlungen abgeleitet werden können. Beispielsweise geht mit einer angemessenen

Geschwindigkeit des Rücklaufs, einer aktiven Lenkungsrückstellung und einem kleinen Restwinkel eine positive Bewertung des Rückstellverhaltens einher.

Tabelle 8. Überblick über die verwendete Methodik der Studien zum Lenkgefühl im Straßenverkehr.

Studie	Stichprobe*	Variation des Lenkgefühls	Fahrsituationen	Art der Bewertung	Anzahl Bewertungskriterien	Bewertungsskala
Wolf (2009)	24 NF	4 Pkw	<ul style="list-style-type: none"> • Autobahn • Landstraße • Stadt • Parkieren 	Absolut	23	Freie Antwort wird in 3-stufige Skala übersetzt
Riedel und Arbinger (1997)	<ul style="list-style-type: none"> • 40 NF • 12 EXP 	6 Varianten in einem Pkw <ul style="list-style-type: none"> • Reifenluftdruck • Beladung etc. 	<ul style="list-style-type: none"> • Landstraße • Doppelter Spurwechsel (ISO/TR 8725:1988) 	Absolut	7	<ul style="list-style-type: none"> • 4-stufig • Alle Stufen benannt

* EXP: Experten / NF: Normalfahrer

Im Gegensatz hierzu liegen bei Riedel und Arbinger (1997) quantitative Zusammenhänge zwischen objektiven und subjektiven Parametern vor. Auf Basis von 15 objektiven Variablen konnten die Bewertungen des Lenkgefühls von Experten auf der Landstraße bis zu 50 % vorhergesagt werden. Für die Normalfahrer hingegen lag die Varianzaufklärung um 10 %. Insgesamt wurden besonders viele Zusammenhänge zwischen den subjektiven Lenkungsbeurteilungen und dem Schwimmwinkel aufgedeckt. Die allgemein niedrigen Varianzaufklärungen werfen jedoch Zweifel auf, ob der verwendete Ansatz zur Erklärung des Lenkgefühls geeignet ist.

Aus diesen zwei sehr unterschiedlichen Studien im Straßenverkehr konnten grundsätzliche Erkenntnisse zum Lenkgefühl gewonnen werden. Diese waren jedoch bezüglich ihrer Aussagekraft begrenzt und lassen nur in einem begrenzten Maße konkrete Rückschlüsse auf die Zusammensetzung des Lenkgefühls zu.

2.4.2. Studien auf dem Testgelände

Nachfolgend werden Studien vorgestellt, welche auf einem Testgelände stattfanden. Diese nutzten u. a. realitätsnahe Fahrsituationen, bei denen das zu absolvierende Streckenprofil typische Fahrsituationen im Straßenverkehr repräsentieren sollte. Dazu wurden beispielsweise auf der Teststrecke vorhandene Stadt- oder Landstraßenparcours befahren (Zschocke, 2009). Ferner wurden synthetische Fahrmanöver genutzt, welche speziell für die Erprobung des Fahr- und Lenkverhaltens konzipiert wurden (Decker, 2009). Diese können nicht unbedingt Fahrsituationen im realen Fahrbetrieb zugeordnet werden. Meist werden die Manöver mit Pylonen auf Freiflächen entsprechend der benötigten Geschwindigkeiten und Querschleunigungen aufgebaut. Ein Überblick über die verwendeten Methoden ist in Tabelle 9 zu finden.

Tabelle 9. Überblick über die verwendete Methodik der Studien zum Lenkgefühl auf dem Testgelände.

Studie	Stich- probe*	Variation des Lenkgefühls	Versuchs- umgebung	Art der Bewertung	Anzahl Bewertungs- kriterien	Bewertungsskala
Anand (2014)	• 14 NF	10 Varianten in einem Pkw • Lenkunterstützung • Steifigkeit etc.	Stadtparcours	Relativ im Vergleich zu Referenz	9	• 10-stufig • Pole und Mitte benannt
Rothhämel (2013)	• 9 NF • 6 S	• 3 Lkw	Manöver auf glatter Oberfläche	Absolut	32, zu 5 Skalen zusammen- gefasst	• Niveau & Gefallen • 7-stufig, Pole und Mitte benannt
Decker (2009)	Nicht bekannt	Versch. Anzahl Varianten in einem Pkw • Übersetzung • Verzögerung • Lenkunterstützung	• Doppelter Spurwechsel (ISO 3888- 2:2011) • Anlenken auf der Geraden	Relativ im Vergleich zu Referenz	8 / 12 pro Abschnitt	• Niveau & Gefallen • 5-stufig, Pole benannt
Zschocke (2009)	• 5 NF • 17 EXP	• 10 Pkw	• Autobahn • Landstraße • Stadt • Doppelter Spurwechsel (ISO 3888- 1:1999)	Absolut	10-14 pro Abschnitt	• Niveau: 6- stufig, Pole benannt • Gefallen: 10- stufig, teilweise benannt
Ackert (2008)	42 NF	6 Varianten in einem Pkw • Lenkübersetzung • Lenkunterstützung	• einfacher Spurwechsel • Wedeltest • Kurven- elemente	Relativ in Paarver- gleichen	17, zu 3 Skalen zusammen- gefasst	• 5-stufig • Alle Stufen benannt
Barthenheier (2004)	197 NF	9 Varianten in einem Pkw • Rückstellmoment • Dämpfung • Reibung	• Autobahn • Landstraße • Stadtparcours	Relativ in Paarver- gleichen	4	• 5-stufig • Alle Stufen benannt

* EXP: Experten / NF: Normalfahrer / S: Sonstige

Wie Riedel und Arbinger (1997) hatte auch Zschocke (2009) zum Ziel, das subjektive Lenkgefühl für die Stichproben der Normalfahrer und Experten auf Basis objektiver Kennwerte vorherzusagen. Für die Experten ergaben sich in allen Fahrsituationen enge Zusammenhänge. Mit multiplen Regressionen konnte beispielsweise die subjektive Bewertung des Anlenk- und Haltemoments zu 57 % bis 75 % vorhergesagt werden. Diese Varianzaufklärungen sind deutlich höher als die aus dem Straßenverkehr (Riedel & Arbinger, 1997). Bei den

Beurteilungen der Normalfahrer ergaben sich jedoch nur für das Wankverhalten signifikante Unterschiede zwischen den Fahrzeugen, sodass auf weitere Analysen verzichtet wurde. Es muss jedoch hinterfragt werden, ob der verwendete Fragebogen und ISO-Spurwechsel für Normalfahrer geeignet waren. So traten beispielsweise bei der Lenkpräzision große Differenzen zwischen der Bewertung desselben Fahrzeugs in der Referenzfahrt zu Beginn und der Wiederholungsfahrt am Ende des Versuchs auf. Hieraus leitete der Autor ab, dass sich die Subjektivbewertungen mit zunehmender Vertrautheit mit dem für Normalfahrer anspruchsvollen Fahrmanöver veränderten. Ferner zeigten verschiedene Studien, dass mit den objektiven Parametern hauptsächlich Aspekte zusammenhängen, welche auf Wahrnehmungen direkt am Lenkrad basieren. Hierzu zählt beispielsweise die Lenkaktivität, die sich in den benötigten Lenkradwinkeln und Lenkkorrekturen widerspiegelt, sowie die haptische Rückmeldung durch das Lenkradmoment (Ackert, 2008). Decker (2009) fand für das Anlenken auf der Geraden besonders für Kennwerte, welche auf dem Verhältnis zwischen Lenkmoment und Lenkwinkel basieren, hohe Zusammenhänge mit den Subjektivbewertungen. Diese Informationen entsprechen dem Lenkgefühl im engeren Sinn (Braess, 2001, 2004). Andererseits liegen Ergebnisse vor, wonach die Variation der Übersetzung zwar zu einer Veränderung im messbaren Lenkwinkelbedarf führt, sich aber nicht in entsprechend unterschiedlichen Bewertungen der subjektiven Direktheit widerspiegelt (Anand, 2014). Demgegenüber wurde ein enger Zusammenhang ($r_{\max} = .90$) zwischen dem objektiv benötigten Lenkwinkelbedarf und der korrespondierenden subjektiven Bewertung gefunden (Decker, 2009). Den Studien ist jedoch gemein, dass für das Lenkmoment bzw. den Lenkmomentverlauf insgesamt enge Zusammenhänge zwischen subjektiven und objektiven Kennwerten gefunden wurden (Anand, 2014; Decker, 2009). Genaueres Augenmerk wurde außerdem auf das subjektive Sicherheitsempfinden gelegt. So hängt die allgemeine Bevorzugung einer Lenkungsvariante stark mit der empfundenen Sicherheit zusammen (Barthenheier, 2004). Ferner ist diese bei Varianten mit direkter Übersetzung und hohem aufzuwendenden Lenkmoment am höchsten ausgeprägt (Ackert, 2008).

Neben der Objektivierung des Lenkgefühls wurde zum Einfluss demografischer Variablen auf das Lenkgefühl geforscht. Bezüglich des Alters der Fahrer wurde festgestellt, dass jüngere Probanden eine niedrigere Dämpfung bevorzugten im Vergleich zu älteren Fahrern (Barthenheier, 2004). Ferner sprachen Frauen niedrigeren aufzuwendenden Lenkmomenten eine höhere Sportlichkeit zu, wohingegen Männer höhere Lenkmomente als sportlicher empfanden. Schlussendlich wurde auf Basis der gesammelten Daten eine Methode entwickelt, die Teilkollektive mithilfe fahrdynamischer Kenngrößen zu unterscheiden. Mithilfe dieser Erkenntnisse kann beispielsweise eine an den Fahrerwunsch anpassbare Lenkunterstützung realisiert werden. Demgegenüber fand Ackert (2008) keine alters- oder geschlechterspezifischen Unterschiede in den Bewertungen des Lenkgefühls. Auch war keine Differenzierung des Probandenkollektivs anhand dieser demografischen Variablen möglich.

In allen diesen Studien wurde die Variation des Lenkgefühls in einem Fahrzeug dargestellt. Diesen Ergebnissen stehen jedoch auch Studien mit verschiedenen Fahrzeugen gegenüber, in denen die Probanden keine Unterschiede zwischen den Fahrzeuglenkungen wahrnahmen (Rothhämel, 2013). Infolgedessen war auch die Berechnung von Zusammenhängen obsolet. Dies könnte u. a. in der unterschiedlichen Erfahrung der Probanden in der subjektiven Bewertung oder ihrer Voreingenommenheit bezüglich der Versuchsfahrzeuge begründet sein.

Zusammenfassend wurden mit Fahrversuchen in realitätsnahen und synthetischen Manövern verschiedene Erkenntnisse zum Lenkgefühl gewonnen.

2.4.3. Studien im Fahrsimulator

Mit der stetigen technischen Weiterentwicklung von Fahrsimulatoren werden diese zukünftig immer mehr an Bedeutung gewinnen. Bertollini und Hogan zeigten bereits 1999, dass mit einem dynamischen Fahrsimulator Aussagen zum Lenkgefühl aus Kundensicht generiert werden können. Ferner hoben sie die Möglichkeit hervor, neuartige Lenksysteme wie beispielsweise Steer-by-Wire-Systeme (Anand, 2014) zu nutzen. Steer-by-Wire-Systeme zeichnen sich dadurch aus, dass der Lenkbefehl elektrisch vom Lenkrad an einen Aktor weitergeleitet wird, der den Lenkbefehl an den gelenkten Rädern ausführt (Harrer & Pfeffer, 2013). Dadurch, dass keine mechanische Verbindung zwischen Lenkrad und gelenkten Rädern besteht, sind noch mehr Gestaltungsmöglichkeiten bei der Lenkungsauslegung vorhanden. Diese machen sich einige der nachfolgend vorgestellten Arbeiten zunutze. Ein Überblick über die verwendeten Methoden ist in Tabelle 10 zu finden.

Einer der Vorteile des Fahrsimulators ist das große Maß an Freiheit, welches bei der Darstellung von Varianten genutzt werden kann. Nichtsdestotrotz bevorzugten die Probanden Lenkradmoment-Varianten, die denen in ihren eigenen Fahrzeugen ähnelten (Anand, 2014). Im Gegensatz zu den Befunden von Barthenheier (2004) liegen aus dem Fahrsimulator Ergebnisse vor, wonach die Präferenzen bezüglich des Lenkradmoments von der Fahrgeschwindigkeit abhängen (Anand, 2014). Diese Präferenzen werden durch die Aspekte Komfort und Kontrollierbarkeit maßgeblich beeinflusst.

Wie auch für die Studien auf dem Testgelände sind konkrete Ergebnisse zum Zusammenhang zwischen den objektiven Parametern und den Subjektivbewertungen zu finden. Auf Basis von bis zu vier objektiven Kenngrößen konnten Varianzaufklärungen bis über 90 % (Dang et al., 2015) erzielt werden. Neben diesen Ergebnissen zum Pkw liegen auch Erkenntnisse zum Lkw vor. Hier wurden u. a. Zusammenhänge zwischen dem objektiv messbaren aufgewendeten Lenkmoment und dem subjektiv wahrgenommenen Lenkmoment sowie dem Rücklauf gefunden (Rothhämel et al., 2011b). Die Analysen mit multiplen Regressionen und neuronalen Netzen führten hierbei zu ähnlichen Ergebnissen.

Weiterhin wird Forschung zum Einfluss der demographischen Daten betrieben, wobei diese über die bereits untersuchten Variablen Alter und Geschlecht (Ackert, 2008; Barthenheier, 2004) hinausgeht. So deckte Black (2010) einige Unterschiede zwischen Fahrerotypen auf, die auch für das Lenkgefühl am Nutzfahrzeug interessant sind. Ein Fahrer-Typ waren bspw. „nutzenorientierte Personen“, die ihr Fahrzeug als Werkzeug ansehen. Da vermutlich die meisten Lkw-Fahrer in diese Gruppe fallen, sind die Ergebnisse für diese Arbeit besonders relevant. Bei diesen Fahrern gingen höhere Lenkmomente mit einer höheren Gefallen-Bewertung einher, während der Zusammenhang bei den anderen Fahrer-Typen umgekehrt war. Aus der Summe dieser Erkenntnisse ergibt sich die Forderung, dass Lenkungsauslegungen nicht nur geschwindigkeitsabhängig, sondern auch individualisierbar sein sollten.

Tabelle 10. Überblick über die verwendete Methodik der Studien zum Lenkgefühl im Fahrsimulator.

Studie	Stich- probe*	Variation des Lenkgefühls	Versuchs- umgebung	Art der Bewertung	Anzahl Bewertungs- kriterien	Bewertungs- skala
Dang et al. (2015)	15 NF	20 Varianten der Lenkunterstützun- g	<ul style="list-style-type: none"> • Statischer Fahrsimulator • Spurwechsel 	Absolut	7	<ul style="list-style-type: none"> • 9-stufig • Alle Stufen benannt
Mandhata et al. (2012)	24 NF	4 Varianten <ul style="list-style-type: none"> • Steifigkeit • Dämpfung usw. 	<ul style="list-style-type: none"> • Statischer Fahrsimulator • Landstraße 	Absolut	9	<ul style="list-style-type: none"> • 7-stufig • Pole benannt
Anand (2014)	30 NF	6 Varianten der Lenkunterstützun- g	<ul style="list-style-type: none"> • Statischer Fahrsimulator • Stadt 	Relativ in Paarver- gleichen	27	<ul style="list-style-type: none"> • 10-stufig • Alle Stufen benannt
Rothhämel et al. (2011b)	28 S	16 Varianten <ul style="list-style-type: none"> • Dämpfung • Trägheit usw. 	<ul style="list-style-type: none"> • Dynamischer Fahrsimulator • Spurwechsel • Kurven- elemente • Slalom 	Relativ im Vergleich zu Referenz	32 Kriterien, zu 5 Skalen zusammen- gefasst	<ul style="list-style-type: none"> • 7-stufig • Pole und Mitte benannt
Black (2010)	43 NF	5 Varianten <ul style="list-style-type: none"> • Lenkunter- stützung • Übersetzung 	<ul style="list-style-type: none"> • Dynamischer Fahrsimulator • Autobahn • Landstraße • Stadt 	Absolut	9	<ul style="list-style-type: none"> • 7-stufig • Pole benannt

* NF: Normalfahrer / S: Sonstige

Zusammenfassend gingen aus diesen Studien auch einige Limitationen hervor, welche hauptsächlich auf die Verständlichkeit des Fragebogens zurückzuführen waren. Dies spiegelt sich beispielsweise darin wider, dass einige Bewertungskriterien von den Probanden nicht verstanden wurden und deshalb keine Analysen möglich waren (Dang et al., 2014; Mandhata et al., 2012). Nichtsdestotrotz trugen die Studien am Fahrsimulator dazu bei, das Lenkgefühl näher zu verstehen und Fortschritte hinsichtlich der Individualisierbarkeit des Lenkverhaltens und Lenkgefühls zu machen.

2.5. Ableitung des Forschungsbedarfs

Die im vorherigen Kapitel vorgestellten Studien zum Lenkgefühl stellen die Basis dieser Arbeit dar. Nachfolgend wird der Forschungsbedarf, welcher bezüglich der subjektiven Bewertung des Lenkgefühls am schweren Nutzfahrzeug besteht, herausgearbeitet.

In Kapitel 2.3 wurden die bisher verwendeten Methoden zur subjektiven Bewertung des Lenkgefühls vorgestellt. Im Folgenden wird herausgestellt, an welchen Punkten die vorliegende Arbeit ansetzt.

Zunächst ist festzustellen, dass nahezu alle Fragebögen zum Lenkgefühl auf den Pkw ausgerichtet sind. Dort ist es von Bedeutung, ein für den Hersteller bzw. das Fahrzeugsegment spezifisches Lenkgefühl zu generieren. Es besteht jedoch stets der Konflikt zwischen einer komfortablen oder sportlichen Lenkungsauslegung (Harrer & Pfeffer, 2013). Im Lkw-Bereich hingegen spielt die Sportlichkeit eine untergeordnete Rolle. Vielmehr ist ein hohes Maß an Komfort in allen Geschwindigkeitsbereichen entscheidend. Da Pkw und Lkw hinsichtlich des Lenkgefühls somit unterschiedliche Anforderungen aufweisen, ist eine Nutzung der Fragebögen aus dem Pkw-Bereich für die Beurteilung des Lkw nicht anzuraten.

Weiterhin zeigte sich bei der Bewertung des Lenkgefühls durch Normalfahrer, dass die in der Praxis häufig verwendeten Experten-Fragebögen für sie nicht geeignet sind. Dies liegt darin begründet, dass sie Fachbegriffe zur Beschreibung der Querdynamik verwenden. Deshalb sind die zu bewertenden Aspekte für Normalfahrer ohne besondere Expertise nicht verständlich und auch die dargestellten Unterschiede nicht wahrnehmbar.

Außerdem sind die Fragebögen speziell für Normalfahrer hinsichtlich ihrer Länge, der verwendeten Kriterien, der Skala etc. stark unterschiedlich. Wird ein Fragebogen mit dem Fokus auf Normalfahrer entwickelt, werden Sachverhalte vereinfacht, um sie an die Expertise der Beurteiler anzupassen. Dies führt zu oberflächlicheren Bewertungen. Hier sind beispielsweise die Fragebögen von Riedel und Arbinger (1997) oder Rothhämel (2013) zu nennen, die aufgrund der geringen Kriterienanzahl bzw. den allgemeinen Fragestellungen keine detaillierten Aussagen zulassen. Eine große Anzahl an Bewertungskriterien (Wolf, 2009; Zschocke, 2009) bringt insbesondere bei in der Bewertung unerfahrenen Probanden die Gefahr der Ermüdung mit sich. Die Detailliertheit einer Bewertung wird auch durch die genutzte Bewertungsskala bestimmt. Während bspw. Wolf (2009) offene Antwortformate verwendet, werden in anderen Studien 10-stufige Antwortformate genutzt (Zschocke, 2009). Bei einigen der gefundenen Kriterien ist außerdem anzuzweifeln, ob sie den Forderungen der psychologischen Methodenlehre hinsichtlich einer optimalen Fragenformulierung gerecht werden können. So soll pro Bewertungskriterium jeweils nur ein Aspekt erfasst werden (Kallus, 2010). Außerdem sollen Items so klar und eindeutig wie möglich formuliert sein. In einigen Fragebögen, z. B. von Wolf (2009) und Rothhämel (2013), wird diesen Forderungen nicht nachgekommen. Ferner fand nur bei wenigen der genannten Arbeiten, z. B. bei Riedel und Arbinger (1997) und Rothhämel (2013), eine umfassende, systematische Evaluation des Fragebogens statt. Diese Überprüfung ist jedoch nötig, bevor der Fragebogen überhaupt zur Generierung von Forschungsergebnissen herangezogen werden sollte.

Die Zusammenstellung verdeutlicht, dass bisher kein etablierter Fragebogen für Normalfahrer vorliegt. Daraus ergibt sich die Notwendigkeit, einen speziell für Normalfahrer geeigneten Fragebogen zur Bewertung des Lenkgefühls am schweren Nutzfahrzeug zu entwickeln und zu validieren. Neben der Entwicklung des Fragebogens selbst ist auch der Versuchsablauf festzulegen. Dieser gibt die nötigen Randbedingungen vor. Somit sind beispielsweise Anforderungen an die Stichprobe, die Versuchsumgebung und das Bewertungsvorgehen abzuleiten. Hieraus folgt:

Forschungsfrage 1: Mit welchem Versuchskonzept können Normalfahrer das Lenkgefühl am schweren Nutzfahrzeug bewerten?

Der Überblick über bisherige Fahrversuche zeigte, dass der öffentliche Straßenverkehr, die Teststrecke und der Fahrsimulator als Versuchsumgebung genutzt werden können (Kapitel 2.3.2). Die Validität der Ergebnisse, welche aus Evaluationen im Straßenverkehr hervorgehen, ist am höchsten im Vergleich zu den anderen Versuchsumgebungen. Dies ist darin begründet, dass die Versuchsumgebung dem alltäglichen Fahrumfeld entspricht. Die Freiheitsgrade in der Variation des Lenkgefühls sind hier jedoch begrenzt, da die Versuchsfahrzeuge eine Straßenfreigabe benötigen. Deswegen werden bei Studien im Straßenverkehr hauptsächlich Hardware-Komponenten variiert oder verschiedene Fahrzeuge genutzt. Zusammenfassend lässt sich also sagen, dass die Erprobung verschiedener Fahrzeuge im öffentlichen Straßenverkehr vergleichsweise einfach umsetzbar ist. Dennoch liegen nur einige Erkenntnisse aus dem Pkw-Bereich vor (Kapitel 2.4).

Einerseits zeigte Wolf (2009), dass Normalfahrer das Lenkgefühl verbal beschreiben können und lenkungsbezogene Unterschiede zwischen den Testfahrzeugen wahrnehmen. Sie können diese jedoch nicht mit den zugrunde liegenden technischen Unterschieden in Bezug bringen. Aufgrund des verwendeten Versuchs- und Auswertungskonzepts liegen aus dieser Studie jedoch nur qualitative Zusammenhänge vor. Im Gegensatz hierzu konnten Riedel und Arbinger (1997) Rückschlüsse von objektiven Kennwerten auf Subjektivurteile ziehen. Hierzu variierten sie Hardware-Komponenten eines Testfahrzeugs und erprobten diese auf einem Landstraßenparcours und im doppelten Spurwechsel. Wie jedoch angesprochen, lagen für das verwendete Fahrmanöver engere Zusammenhänge zwischen den Kennwerten vor als für die Landstraße.

Obwohl in beiden Studien Erkenntnisse zum weiteren Verständnis des Lenkgefühls gewonnen wurden, blieben diese in ihrer Detailliertheit und Belastbarkeit hinter denen aus den anderen Versuchsumgebungen zurück. Somit wurden auch die Grenzen eines derartigen Versuchsdesigns deutlich. Unklar ist jedoch, inwiefern diese Grenzen durch eine Optimierung der verwendeten Methodik ausgeweitet werden können. Deshalb soll diese Herangehensweise weiter entwickelt, auf den Lkw übertragen und hinsichtlich ihrer Eignung bewertet werden. Hieraus folgt:

Forschungsfrage 2: Kann das Lenkgefühl verschiedener Fahrzeuge unter Anwendung eines geeigneten Versuchskonzepts in realitätsnahen Fahrversuchen evaluiert werden?

In Kapitel 2.4 wurden verschiedene Studien zur Objektivierung des Lenkgefühls, also der Vorhersage der subjektiven Bewertungen aus den objektiven Kennwerten, vorgestellt.

Aus einigen dieser Studien, welche realitätsnahe Fahrmanöver nutzten, gingen fundierte Ergebnisse hervor. Barthenheier (2004) stellte in einem Fahrzeug mehrere Varianten des Lenkradmoments dar und lies diese in einem Autobahn-, Landstraßen- und Stadtparcours bewerten. Die gewonnenen Ergebnisse wiesen Einflüsse des Alters und des Geschlechts auf die Subjektivbewertungen nach. In einer anderen Studie wurden verschiedene Varianten des Lenkgefühls in einem Stadtparcours und bei höheren Geschwindigkeiten evaluiert (Anand, 2014). Es wurden Zusammenhänge zwischen der Lenkunterstützung und den subjektiven Bewertungen aufgedeckt. Eine besondere Bedeutung für das Lenkgefühl kam dabei dem wahrgenommenen Lenkmoment zu. Lediglich bei Rothhämel (2013) führte die Bewertung von drei verschiedenen Lkw in verschiedenen Fahrmanövern nicht zu belastbaren Ergebnissen. So wurden von der Versuchsgruppe keine Unterschiede zwischen den Testfahrzeugen festgestellt.

Dies wird im Wesentlichen auf die Nutzung verschiedener Testfahrzeuge und die damit verbundene Beeinflussung der Testpersonen zurückgeführt.

Andererseits wurden auch Studien durchgeführt, welche auf der Nutzung synthetischer Manöver basierten. Ackert (2008) beispielsweise nutzte den einfachen Spurwechsel, Wedeltest und Kurvenelemente, um sechs Lenkungsvarianten evaluieren zu lassen. Es wurden vor allem Zusammenhänge zwischen objektiven Kennwerten und subjektiven Bewertungen von Parametern, welche direkt am Lenkrad wahrnehmbar sind, gefunden. Unter Nutzung einer ähnlichen Vorgehensweise wurde ein Zusammenhang der Subjektivbewertungen mit der Querschleunigung und dem Lenkmoment gefunden (Decker, 2009). Die Bewertung fand hier im doppelten Spurwechsel statt. Beim Anlenken auf der Geraden hingen die Bewertungen mit dem Verhältnis zwischen Lenkmoment und Lenkwinkel zusammen. Weniger aussagekräftige Ergebnisse erzielte Zschocke (2009) mit seiner Teilstichprobe der Normalfahrer. Er nutzte einen straßenverkehrsähnlichen Parcours und den Spurwechsel, um verschiedene Fahrzeuge von Normalfahrern evaluieren zu lassen. Die Beurteilungen der Normalfahrer wiesen lediglich in einigen Aspekten Unterschiede auf. Dies war vermutlich darauf zurückzuführen, dass Fragebogen, Fahrmanöver und Parcours sowie die gesamte Evaluationsmethodik auf Experten als Beurteiler ausgelegt waren.

Die vorliegenden Ergebnisse aus Studien auf dem Prüfgelände legen insgesamt nahe, dass diese Versuchsumgebung geeignet ist, um Daten zur Objektivierung des Lenkgefühls zu sammeln. Die beschriebenen Einschränkungen unterstreichen aber die Notwendigkeit nach einem geeigneten Versuchskonzept. Deshalb soll die zu Forschungsfrage 1 entwickelte Methode zur Bewertung des Lenkgefühls von Normalfahrern genutzt werden. Hieraus folgt:

Forschungsfrage 3: Können unter Anwendung eines geeigneten Versuchskonzepts Erkenntnisse zur Objektivierung des Lenkgefühls aus Sicht der Normalfahrer generiert werden?

Wie zuvor diskutiert liegen widersprüchliche Erkenntnisse darüber vor, ob Normalfahrer zuverlässige Bewertungen des Lenkgefühls abgeben können (siehe Kapitel 2.3.1).

Für die Zuverlässigkeit der Bewertungen spricht beispielsweise das Ergebnis, dass Probanden Varianten des Lenkgefühls bei wiederholter Bewertung nicht signifikant verschieden einschätzten (Anand, 2014). Andere Studien belegen, dass die auf Bewertungen von Normalfahrern basierenden Varianzaufklärungen vergleichbar hoch bzw. sogar höher sind als die, die auf Experten-Evaluationen basieren (Dang et al., 2015). Dieses Ergebnis ist besonders interessant, da die Bewertungen der Experten in einem dynamischen, die der Normalfahrer aber in einem statischen Fahrsimulator erfolgten. Somit konnten trotz der weniger realistischen Simulation entsprechende Ergebnisse erzielt werden (Zöller, 2015).

Andererseits gibt es Belege, dass mit Normalfahrern keine zuverlässigen Ergebnisse erzielt werden können. So zeigte Zschocke (2009), dass Normalfahrer nur bestimmte Parameter zuverlässig bewerten können. Zum einen nahmen die Probanden nur bezüglich des Wankverhaltens signifikante Unterschiede zwischen den Fahrzeugen wahr. Ferner waren die Abweichungen zwischen Referenz- und Wiederholungsfahrt bei den Normalfahrern tendenziell höher als bei den Experten. Unterschiede im Antwortverhalten zwischen Normalfahrern und Experten zeigten sich auch in einem Fahrversuch zur subjektiven Bewertung des gesamten Fahrverhaltens (Schoeggel &

Ramschak, 2000). Hier waren die Interquartilsabstände bei Kunden tendenziell größer als bei Experten. Hieraus leiteten die Autoren ab, dass Kunden in der Lage sind, Probleme exakt zu erkennen, nicht aber sie fein in Bezug auf einzelne Kriterien zu bewerten.

Es ist also unklar, ob Normalfahrer in der Lage sind, zuverlässige Subjektivbewertungen abzugeben. Hieraus folgt:

Forschungsfrage 4: Wie können Normalfahrer das Lenkgefühl am schweren Nutzfahrzeug zuverlässig bewerten?

Der Arbeit lagen somit vier Forschungsfragen zugrunde. Diese sind zusammen mit den durchgeführten Forschungsarbeiten in Tabelle 11 zu finden. Das Vorgehen gliederte sich in die folgenden Schritte auf: Zunächst wurde ein erster Entwurf des Versuchskonzepts unter Einbeziehung von Fahrern entwickelt. Dieses wurde anschließend in einem Fahrversuch im öffentlichen Straßenverkehr mit verschiedenen Fahrzeugen evaluiert. Infolgedessen wurde eine erneute Anpassung des Konzepts vorgenommen. Das optimierte Versuchskonzept diente dann erneut der Evaluation verschiedener Fahrzeuge im öffentlichen Straßenverkehr. Nach der anschließenden Finalisierung wurde es im dritten Fahrversuch eingesetzt. Auf Basis des Vergleichs verschiedener Varianten der Lenkunterstützung in Fahrmanövern wurden Erkenntnisse zum Lenkgefühl und zu dessen Optimierung gewonnen.

Tabelle 11. Forschungsfragen und dazugehörige Forschungsarbeiten.

Forschungsfrage	Forschungsarbeit
1 Mit welchem Versuchskonzept können Normalfahrer das Lenkgefühl am schweren Nutzfahrzeug bewerten?	<ul style="list-style-type: none">• Entwicklung eines Versuchskonzepts (Kapitel 3.1)• Evaluation in Studie I (Kapitel 3.2)• Evaluation in Studie II (Kapitel 3.3)
2 Kann das Lenkgefühl verschiedener Fahrzeuge unter Anwendung eines geeigneten Versuchskonzepts in realitätsnahen Fahrversuchen evaluiert werden?	<ul style="list-style-type: none">• Studie I (Kapitel 3.2)• Studie II (Kapitel 3.3)
3 Können unter Anwendung eines geeigneten Versuchskonzepts Erkenntnisse zur Objektivierung des Lenkgefühls aus Sicht der Normalfahrer generiert werden?	<ul style="list-style-type: none">• Studie III (Kapitel 4)
4 Wie können Normalfahrer das Lenkgefühl am schweren Nutzfahrzeug zuverlässig bewerten?	<ul style="list-style-type: none">• Studie I (Kapitel 3.2)• Studie II (Kapitel 3.3)• Studie III (Kapitel 4)

3. Entwicklung und Evaluation eines Versuchskonzepts zur subjektiven Bewertung des Lenkgefühls durch Normalfahrer

3.1. Entwicklung des Versuchskonzepts

Wie oben aus der Literaturgrundlage abgeleitet, sollte ein Versuchskonzept zur Evaluation des Lenkgefühls durch Normalfahrer entwickelt werden. Unter dem Versuchskonzept wird dabei zum einen der Fragebogen zur Erfassung der subjektiven Bewertungen verstanden. Zum anderen ist damit auch der Versuchsaufbau gemeint, also der Rahmen, in dem der Fragebogen bearbeitet wird. Die Entwicklung erfolgte in elf Schritten, welche in Abbildung 8 dargestellt sind. Diese werden in den folgenden Kapiteln vorgestellt.

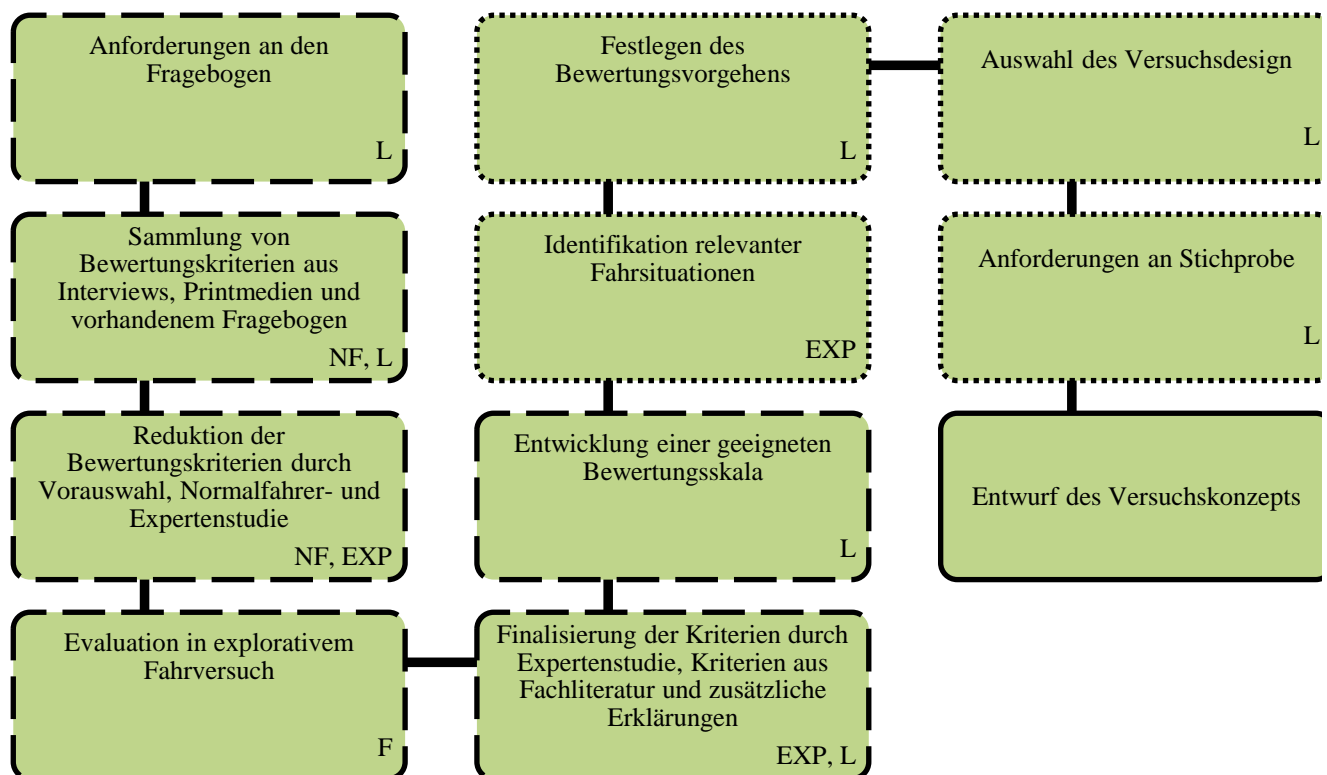


Abbildung 8. Schritte der Entwicklung des Fragebogens (gestrichelt) und der Definition der Anforderungen an den Versuchsaufbau (gepunktet) unter Angabe der verwendeten Quellen Experte (EXP), Fahrer (F), Literatur (L) und Normalfahrer (NF).

3.1.1. Entwicklung des Fragebogens

Anforderungen

Die Anforderungen an den Fragebogen (siehe Kapitel 3.1) sollen nun erläutert und weiter detailliert werden.

- Fragebogen entspricht Expertise von Normalfahrern

Hierunter versteht man zum einen, dass die Bewertungskriterien des Fragebogens insgesamt für Normalfahrer ohne Erfahrung in der Beurteilung des Lenkgefühls geeignet sind. Dies beinhaltet, dass Normalfahrer ein begrenztes Faktenwissen über die technischen Zusammenhänge im Bereich der Lenkung haben. Ferner sind sie

nicht darauf geschult, auf einzelne Aspekte der Lenkung zu achten. Deshalb müssen die verwendeten Bewertungskriterien leicht verständlich sein, was u. a. durch die Vermeidung von Fachbegriffen erreicht werden kann (Gies & Marusic, 2000). Dies gilt in besonderem Maße, da auch eine Beantwortung des Fragebogens im Fahrzeug gefordert wird. Dies schließt ein, dass die Probanden die Kriterien verstehen können, auch wenn sie ihnen nur vorgelesen werden. Zum anderen muss auch die Skala für Normalfahrer verständlich sein. Dies bedeutet, dass der vorgegebene Detaillierungsgrad der Antwortalternativen ihren Fähigkeiten entspricht.

- Bewertungskriterien auf objektive Kenngrößen bezogen

Im Rahmen der Anpassung der Bewertungskriterien an die Expertise der Normalfahrer entsteht häufig ein Konflikt zwischen der Verständlichkeit und Konkretheit der Kriterien. Ein Beispiel hierfür ist die subjektiv empfundene Sicherheit: Während es für den Fahrer vermutlich sehr einfach ist, sein Sicherheitsempfinden zu bewerten, lässt ein solches Bewertungskriterium wenig konkrete Rückschlüsse zu. Es ist also nicht möglich, daraus notwendige Adaptionen in Bezug auf eine bestimmte Kenngröße des Lenksystems abzuleiten. Deshalb ist darauf zu achten, dass Bewertungskriterien leicht verständlich und dennoch so konkret wie möglich auf eine spezifische objektive Kenngröße bezogen sind (Badiru, 2014).

- Alle Aspekte des Lenkgefühls abgedeckt

Es sind einige Studien zu finden, bei denen lediglich bestimmte Aspekte des Lenkgefühls erfasst wurden. Beispiele hierfür ist der Fokus auf das Sicherheitsempfinden bei Ackert (2008) oder auf das Lenkgefühl bei höheren Geschwindigkeiten bei Decker (2009). Der nun zu entwickelnde Fragebogen soll aber das Lenkgefühl vollumfänglich bewertbar machen.

- Bewertungskriterien speziell auf schweres Nutzfahrzeug ausgerichtet

Hierunter versteht sich, dass die für das Lenkgefühl am schweren Nutzfahrzeug relevanten Aspekte gesammelt werden. Beispielsweise sind der Korrekturbedarf bei der Geradeausfahrt oder die Lenkbarkeit bei niedrigen Geschwindigkeiten beim Lkw von besonderer Bedeutung (Shyrokau, Loof, Stroosma, Wang & Happee, 2015). Andererseits sollen keine Aspekte eingeschlossen werden, die für den Lkw nicht relevant sind. Hierunter fällt z. B. ein ausgeprägter Fokus auf die Sportlichkeit der Lenkung.

- Zufriedenstellende psychometrische Qualität

Unter der psychometrischen Qualität wird die Qualität eines Datenerhebungsverfahrens verstanden (Kallus, 2010). Sie wird hauptsächlich über die Objektivität, Reliabilität und Validität der Messung bestimmt (Renner et al., 2012). Nur wenn der Fragebogen gewisse Anforderung in dieser Hinsicht erfüllt, darf er genutzt werden.

Auf Basis dieser fünf Anforderungen wurde nun der Fragebogen entwickelt. Wie bereits in der Motivation dargestellt, soll der Nutzer im Fokus der gesamten Arbeit stehen. Dies bedeutet, dass Lkw-Fahrer in alle wesentlichen Entwicklungsschritte eingebunden werden.

Sammlung von Bewertungskriterien

Entsprechend obiger Anforderungen sollen Bewertungskriterien generiert werden, welche der Expertise von Normalfahrern entsprechen und das Lenkgefühl am Nutzfahrzeug abbilden. Deshalb sollten in einem ersten Schritt die Aspekte, die Fahrer zur Beschreibung des Lenkens und des Lenkgefühls bekannt sind, gesammelt werden. Dieses Vorgehen ist analog dem von Jürgensohn und Kolrep (2006), Riedel und Arbinger (1997) und Rothhämel et al. (2010). Es wurden $N = 31$ Lkw-Fahrer interviewt. Die Fahrer waren im Mittel 50,0 Jahre alt ($SD = 13,0$), fuhren seit 23,6 Jahren Lkw ($SD = 15,2$) und hatten eine jährliche Fahrleistung von 86.419 km ($SD = 46.914$). Alle Teilnehmer waren männlich. Die Befragten stammten aus dem gesamten Bundesgebiet und waren im Verteiler-, Fern- und Baustellenverkehr tätig. Es wurden lediglich Berufskraftfahrer mit guten Deutschkenntnissen interviewt. Die Befragungen fanden auf Rasthöfen statt und dauerten im Durchschnitt weniger als 15 min pro Teilnehmer.

Die Probanden wurden aufgefordert, alle Aspekte zu nennen, mit denen die Lenkung eines Lkw beschrieben werden kann. Hierbei wurde auch der Hinweis gegeben, nicht nur an die Lenkung im derzeit gefahrenen Fahrzeug zu denken, sondern auch an besonders gute oder schlechte Systeme in anderen Fahrzeugen. Im zweiten Schritt wurde gefragt, welche Aspekte in welchen Fahrsituationen am wichtigsten sind. Hier wurden die Situationen Autobahn, Landstraße, Stadt und Rangieren vorgegeben.

Insgesamt wurden 85 Kriterien gesammelt, wobei jeder Proband im Mittel 8,5 Aspekte nannte ($Min = 3$, $Max = 17$, $SD = 3,67$). Einige Aspekte waren sehr spezifisch, z. B. *gute Rückmeldung der Straße* oder *leichtgängig im Stand*. Andere Aspekte waren wiederum sehr allgemein, beispielsweise *ruhig* oder *gefühllos*. Trotz der großen Zahl an Begriffen wurde deutlich, dass die Probanden häufig ein sehr begrenztes Vokabular zur Beschreibung der Lenkung zur Verfügung hatten. Dies wurde seitens der Probanden oftmals damit begründet, dass das Lenken automatisch ablaufe und ihnen nicht bewusst sei.

Im Sinne der geforderten Vollständigkeit des Fragebogens wurden zusätzlich Kriterien aus einschlägigen Printmedien zum Lastkraftwagen gesammelt. Da diese Medien von Berufskraftfahrern gelesen werden, bilden sie ihre Wissensgrundlage zum Thema Lenkung. Somit ist davon auszugehen, dass die genutzten Begrifflichkeiten den Fahrern bekannt und für sie verständlich sind. Beispiele der 25 extrahierten Aspekte sind *stabil gegenüber Schlaglöchern* und *gelassen*.

Schlussendlich sollten auch in der Lenkungsbeurteilung etablierte Kriterien eingeschlossen werden. Wie oben vorgestellt, sind die Fragebögen aus der Literatur für die vorliegende Fragestellung nur bedingt geeignet. Lediglich der Fragebogen von Rothhämel et al. (2010) entsprach aufgrund des Fokus auf Normalfahrer und das schwere Nutzfahrzeug den Anforderungen. Nach einer Übersetzung ins Deutsche wurden alle 35 Kriterien übernommen, darunter *kursstabil* und *komfortabel*.

Insgesamt lagen am Ende der Sammlung 145 Bewertungskriterien vor.

Reduktion der Bewertungskriterien

Im zweiten Schritt wurden Aspekte ausgeschlossen, welche in den drei Quellen mehrfach genannt wurden, beispielsweise *direkt*. Weiterhin wurden solche Kriterien zusammengefasst, die inhaltlich sehr ähnlich sind. Beispielsweise wurden *leicht* und *leichte Übersetzung* zu *leichtgängig* zusammengefasst. Drittens wurden zu spezifische Aspekte exkludiert, z. B. *beim Bremsen darf die Lenkung nicht abhauen*. Schlussendlich wurden unverständliche Kriterien ausgeschlossen, z. B. *behände* und *intuitiv*. Somit verblieben 53 Kriterien.

Aus Gründen der Versuchsökonomie war diese Anzahl an Kriterien jedoch noch zu hoch. Deshalb wurde, in Anlehnung an Rothhämel et al. (2010), eine weitere Reduktion durchgeführt. Um auch hier den zukünftigen Nutzer einzubinden, wurden $N = 45$ Lkw-Fahrer interviewt. Im Mittel waren die Fahrer 46,0 Jahre alt ($SD = 12,8$), fuhren seit 18,8 Jahren Lkw ($SD = 12,9$) und hatten eine jährliche Fahrleistung von 120.000 km ($SD = 50.000$). Alle Fahrer waren männlich. Die Befragten stammten aus dem gesamten Bundesgebiet und waren im Verteiler-, Fern- und Baustellenverkehr tätig. Es wurden lediglich Berufskraftfahrer mit guten Deutsch-kenntnissen zugelassen. Die Befragungen fanden auf Rasthöfen statt und dauerten im Durchschnitt weniger als 10 min pro Teilnehmer.

Die Probanden sollten auf einer dichotomen Skala (ja/nein) angeben, ob die 53 gesammelten Kriterien zur Beschreibung des Lenkgefühls geeignet sind. Sie wurden darauf hingewiesen, dass die Aspekte sowohl gute als auch schlechte Eigenschaften des Lenkgefühls beschreiben können. Es wurde festgelegt, alle Kriterien beizubehalten, die von mindestens 50 % der Fahrer akzeptiert wurden. Die Analyse der Häufigkeiten zeigte, dass dies für 30 Kriterien der Fall war. Aus einer qualitativen Analyse gingen jedoch zwei Problemfelder hervor: Erstens traten bei einigen Kriterien häufiger Verständnisprobleme seitens der Probanden auf. Zweitens waren unter den ausgeschlossenen Kriterien besonders viele, die negative Eigenschaften abdecken. Es ist anzunehmen, dass die Probanden trotz gegenteiliger Instruktion lediglich solche Bewertungskriterien beibehalten wollten, die eine für sie angenehme Lenkung beschreiben.

Um sicherzustellen, dass positive wie negative Kriterien gleichermaßen vertreten sind, wurden $N = 9$ Fahrdynamik-Experten interviewt. Auch sie sollten auf einer dichotomen Skala (ja/nein) angeben, ob die 53 gesammelten Kriterien zur Beschreibung des Lenkgefühls geeignet sind. Auch für sie wurde festgelegt, alle Kriterien beizubehalten, die von mindestens 50 % der Probanden akzeptiert wurden. Die Analyse der Häufigkeiten zeigte, dass 35 Kriterien von mehr als 50 % der Experten akzeptiert wurden. Davon wurden 13 vorher von den Fahrern ausgeschlossen.

Somit wurde festgelegt, dass alle Kriterien beibehalten werden, die von mindestens 50 % der Lkw-Fahrer oder 50 % der Experten akzeptiert werden. Dies war für 43 Kriterien der Fall.

Evaluation in explorativem Fahrversuch

Da sich der Fragebogen noch im Entwicklungsprozess befand, bestand das Ziel der ersten Evaluation in der Überprüfung der Verständlichkeit der Kriterien. In diesem Schritt wurde bewusst nicht auf Normalfahrer zurückgegriffen, weil eine kritische Reflektion nötig war. Vielmehr wurden Personen herangezogen, die

Erfahrung mit der subjektiven Bewertung des Fahrkomforts hatten, jedoch nicht in Bezug auf die Lenkung. So nahmen an dem Fahrversuch $N = 6$ Werkstattmitarbeiter teil. Der Fahrversuch beinhaltete Abschnitte auf einem Lkw-Parkplatz, in der Stadt, auf Landstraßen und der Autobahn und dauerte ca. 1 h pro Person. Es wurden zwei unterschiedliche Sattelzugmaschinen genutzt. Die Probanden wurden während der gesamten Fahrt von einem Versuchsleiter begleitet.

Die Probanden wurden zu Beginn des Versuchs über den Entwicklungsstand des Fragebogens und die Ziele des Versuchs informiert. Sie wurden außerdem vor und während des Versuchs explizit darauf hingewiesen, alle Unklarheiten sowie Verbesserungsvorschläge zu äußern. Ähnlich des später vorgestellten Bewertungsvorgehens (Kapitel 3.1.2) wurden die Bewertungskriterien während der Fahrt vom Versuchsleiter einzeln vorgestellt. Die Teilnehmer sollten sich zuerst in die Rolle eines regulären Probanden versetzen und eine Bewertung des Fahrzeugs vornehmen. Hierzu sollten sie angeben, inwiefern dieses Kriterium auf das gefahrene Fahrzeug zutrifft. Anschließend sollten sie das Kriterium hinsichtlich seiner Verständlichkeit und Relevanz fürs Lenkgefühl einschätzen.

Zum einen meldeten alle Fahrer zurück, dass die derzeitige Anzahl an Kriterien viel zu groß war. Die Fahrer berichteten außerdem, dass einige Kriterien noch zu unspezifisch waren. Beispielsweise konnten sie die Kriterien *geschmeidig* oder *folgsam* nicht auf technische Zusammenhänge beziehen. Ferner war der Unterschied zwischen einzelnen Kriterien nicht klar. So konnten die Probanden z. B. die Kriterien *präzise* und *zielgenau* nicht differenzieren. Andererseits gaben sie an, dass einige wichtige Aspekte des Lenkgefühls bisher noch nicht vertreten waren, bspw. die Lenkungsstößigkeit. Außerdem zeigte die Analyse der Begründungen der Fahrer, dass die Probanden bei einzelnen Kriterien extrem unterschiedliche Begründungen nannten. Dies stärkte die Annahme, dass die Kriterien weiter spezifiziert werden mussten.

Finalisierung der Bewertungskriterien

Zuerst sollte die Anzahl der Kriterien weiter verringert werden. Wie sich jedoch bei der letzten Reduktion zeigte, konnten hierfür keine Normalfahrer genutzt werden. Deshalb wurden $N = 9$ Fahrdynamik-Experten herangezogen. Sie sollten zum einen angeben, welche der 43 Kriterien Synonyme oder Antonyme sind. Insgesamt wurden auf diese Art 16 synonyme Kriterien identifiziert, beispielsweise *reagiert sofort* und *direkt*. Das jeweils im Sprachgebrauch geläufigere Kriterium wurde beibehalten. Bezüglich der zur vollständigen Beschreibung des Lenkgefühls fehlenden Aspekte wurde eine erneute Literaturrecherche durchgeführt. Es wurden acht Kriterien aus bestehenden Fragebögen übernommen (Harrer, 2007; Schalz, Duhr & Marusic, 2002; Wolf, 2009; Zschocke, 2009). Somit bestand der finale Entwurf des Fragebogens aus 35 Kriterien.

Anschließend wurden zu den einzelnen Bewertungskriterien Erklärungen entwickelt, wie sie beispielsweise auch bei Zschocke (2009) und Koch (2010) zu finden sind. Diese sollten die Rückführbarkeit auf technische Zusammenhänge und die Differenzierbarkeit zwischen Kriterien verbessern. Die Erklärungen dienten zur Beschreibung oder Umschreibung des Kriteriums. Auch die Erklärungen durften jedoch keine fachspezifische Terminologie enthalten. Dennoch musste die Erklärung so spezifisch und detailliert wie möglich sein. Für zehn Kriterien wurden keine Erklärungen entwickelt, da diese keinen Mehrwert boten. Mithilfe dieser Erklärungen sollte

umgangen werden, dass, wie beispielsweise bei Rothhämel et al. (2011a) beschrieben, unzureichend genau erklärte Kriterien zu einer erhöhten Streuung der Bewertungen führen. Der erste Fragebogenentwurf ist in Kapitel 3.1.3 zu finden. Eine Auswahl der Bewertungskriterien ist in Tabelle 12 zu finden. Insgesamt lagen 17 invertierte Items vor, z. B. *hat Spiel* und *schwammig*. Diese entgegen dem Konstrukt ausgerichteten Items sind mit (i) gekennzeichnet. Ferner lagen Kriterien ohne eine ergänzende Erklärung vor, beispielsweise *beherrschbar* und *schwammig*.

Tabelle 12. Auswahl der Bewertungskriterien mit dazugehöriger Erklärung.

Bewertungskriterium*	Erklärung
Beherrschbar	/
Hat Spiel (i)	Es gibt einen toten Bereich um die Mitte.
Rückstellverhalten	Die Lenkung stellt sich ausreichend selbsttätig in die Geradeausstellung zurück.
Schwammig (i)	/

* (i): invertiert

Entwicklung der Bewertungsskala

Abschließend wurden die Kriterien mit einer Rating-Skala versehen. Wie in Kapitel 2.3.3 vorgestellt, konnte keine der bei bisherigen Forschungen zum Lenkgefühl verwendeten Skalen übernommen werden. Dennoch wird der vorgestellte Ansatz (Kim, 2011; Kim & Yoon, 2015; Zschocke, 2009), die Bewertung in zwei Schritte aufzuteilen, übernommen.

Im ersten Schritt sollten die Beurteiler somit angeben, inwiefern ein gewisser Aspekt bei der eben gefahrenen Lenkung ausgeprägt ist. Für diese Niveau-Einschätzung fungiert der Mensch im Rahmen seiner Möglichkeiten als objektives Messinstrument. Somit sollten alle Beurteiler zu einer ähnlichen Einschätzung kommen. Im zweiten Schritt sollten die Probanden ihre Gefallen-Bewertung abgeben. Bezugnehmend auf die Niveau-Bewertung war nun anzugeben, inwiefern ihnen der Grad der Ausprägung eines bestimmten Kriteriums gefällt. Diese Einschätzung ist individuell verschieden und hängt stark von den persönlichen Präferenzen der Person ab. Dies bedeutet, dass Fahrer die gleiche Ausprägung des Niveaus hinsichtlich des Gefallens völlig unterschiedlich einschätzen können. Zusätzlich ist anzumerken, dass die Gefallen-Bewertung in Abhängigkeit von der Situation zu sehen ist (Gil Gómez et al., 2015). Im Zuge der Datenerhebung ist es von größter Bedeutung, den Probanden immer wieder auf den Zusammenhang zwischen der Niveau- und der Gefallen-Bewertung hinzuweisen. Schließlich soll vermieden werden, dass die Gefallen-Bewertung global für das gesamte Lenksystem oder das gesamte Fahrzeug abgegeben wird.

Absolute Bewertung

Zeigt das Fahrzeug / die Variante diese Eigenschaft?										
1	2	3	4	5	6	7	8	9	10	
gar nicht										völlig
Gefällt Ihnen, wie stark diese Eigenschaft ausgeprägt ist?										
1	2	3	4	5	6	7	8	9	10	
gar nicht										völlig

Relative Bewertung

Wie stark zeigt das Fahrzeug / die Variante diese Eigenschaft im Vergleich zum vorher bewerteten Fahrzeug / zur vorher bewerteten Variante / zur Referenz?										
-5	-4	-3	-2	-1	0	1	2	3	4	5
sehr viel weniger										sehr viel mehr
Wie sehr gefällt Ihnen, wie stark diese Eigenschaft ausgeprägt ist, im Vergleich zum vorher bewerteten Fahrzeug / zur vorher bewerteten Variante / zur Referenz?										
-5	-4	-3	-2	-1	0	1	2	3	4	5
sehr viel weniger										sehr viel mehr

Abbildung 9. Skalen zur absoluten und relativen Niveau- und Gefallen-Bewertung.

Nach dieser Aufteilung musste die Anzahl der Stufen festgelegt werden. Wie vorher diskutiert, sollen im Sinne der Übersichtlichkeit beide Aspekte mit der gleichen Anzahl an Antwortkategorien versehen werden. Da zu diesem Zeitpunkt noch nicht abgeschätzt werden konnte, wie detailliert die Beurteilungen der Fahrer sind, sollten in keinem Fall zu wenige Kategorien angeboten werden. Nach Heißing und Brandl (2002) setzte sich die 10-stufige Skala in der Automobilindustrie weitgehend durch und wurde auch von Harrer (2007), Kim und Yoon (2015) und Zschocke (2009) verwendet. Auch in der psychologischen Forschung werden 10-stufige Skalen häufig verwendet. Für die Benennung der Skalenpole wurden verschiedene Quellen als Ideengeber herangezogen (Bass, Cascio & O'Connor, 1974; Bühner, 2011; Kallus, 2010; Moosbrugger & Kelava, 2012). Bei der Skala zur

absoluten Bewertung wurde dem Minimum der Zahlenwert 1 und die Beschreibung *gar nicht* zugewiesen, während das Maximum den Zahlenwert 10 und die Beschreibung *völlig* erhielt (Abbildung 9). Für die relative Bewertung in Bezug auf eine Referenz, eine Variante oder ein Fahrzeug wird eine 11-stufige Skala verwendet. Diese bietet einen zusätzlichen Nullpunkt, der für keinen wahrgenommenen Unterschied steht. Das Minimum stellt der Zahlenwert -5 mit der Beschreibung *sehr viel weniger* dar, das Maximum der Zahlenwert +5 mit der Beschreibung *sehr viel mehr*.

3.1.2. Versuchsaufbau

Fahrsituationen

Die Bewertung des Lenkgefühls soll während dem Durchfahren relevanter Fahrsituationen (FS) erfolgen, da die Eindrücke nur so unmittelbar und unbeeinflusst gesammelt werden können (Wolf, 2009). Hierzu ist es nötig, für jedes Kriterium eine zur Beurteilung geeignete Situation zu definieren. Ferner sind Aspekte der Versuchswirtschaft zu beachten.

Den entwickelten Bewertungskriterien sollten nun geeignete Fahrsituationen zugeordnet werden, in welchen die zu bewertenden Aspekte von den Probanden erlebt werden können. Es ergaben sich folgende Anforderungen (Ackert, 2008; Rothhämel, 2013; Rothhämel et al., 2011a):

- FS_1: Die gewählten Fahrsituationen bilden den gesamten Fahrbetrieb ab, also z. B. niedrige und hohe Fahrzeuggeschwindigkeiten sowie kleine und große Lenkwinkel.
- FS_2: Die gewählten Fahrsituationen entsprechen dem fahrerischen Können der Normalfahrer. Im Straßenverkehr zieht dies eine Anpassung der möglicherweise vom Versuchsleiter vorgegebenen Fahrgeschwindigkeiten mit sich. Bei Fahrmanövern auf dem Prüfgelände ist die Beherrschbarkeit des Fahrzeugs sicherzustellen. Die Querschleunigung kann hierbei als Orientierung dienen. Während im Pkw maximale Querschleunigungen von mehr als 7 m/s^2 auftreten (Mansfield & Whiting-Lewis, 2004), sind diese im Lkw deutlich geringer. Auch im Fahrsimulator ist darauf zu achten, dass die dargestellten Fahrsituationen von den Probanden beherrscht werden können. Obwohl hier kein Gefährdungspotential vorliegt, können zu schwierige Fahrmanöver Auswirkungen auf die Motivation des Probanden haben.
- FS_3: Die Auswahl und Anordnung der Fahrsituationen führt zu einer möglichst kurzen Fahrzeit, um bspw. Ermüdung zu vermindern.
- FS_4: Für jedes Bewertungskriterium liegt, sofern möglich, eine genaue Spezifizierung der zugehörigen Fahrsituationen vor.

Aus der letzten Anforderung ergab sich die Notwendigkeit, diese Fahrsituationen für die vorliegenden 35 Bewertungskriterien zu definieren. Mithilfe eines Workshops mit $N = 8$ Experten wurden für die meisten Kriterien spezifische Situationen identifiziert. Wie in Tabelle 13 dargestellt, kann beispielsweise das Kriterium *Rückstellverhalten* besonders im niedrigen Geschwindigkeitsbereich nach großen benötigten Lenkwinkeln erlebt werden. Andere Kriterien decken jedoch globalere Aspekte des Lenkgefühls ab und sind deshalb nicht an bestimmte Situationen gebunden, bspw. *schwammig*.

Tabelle 13. Auswahl der Bewertungskriterien mit dazugehöriger Erklärung und Fahrsituation.

Bewertungskriterium*	Erklärung	Spezifische Fahrsituation
Beherrschbar	/	Engstelle
Hat Spiel (i)	Es gibt einen toten Bereich um die Mitte.	Geradeausfahrt bei hohen Geschwindigkeiten
Rückstellverhalten	Die Lenkung stellt sich ausreichend selbsttätig in die Geradeausstellung zurück.	Nach großen Lenkwinkeln bei niedrigen Geschwindigkeiten
Schwammig (i)	/	Global

* (i): invertiert

Bewertungsvorgehen

Die Vorgehensweise bei der Beantwortung des Fragebogens soll eine möglichst fundierte Bewertung ermöglichen. Demnach wurde entschieden, in der ersten Evaluation des Fragebogens alle Aspekte einzeln erlebbar zu machen. Darauf aufbauend wurde das Bewertungsvorgehen (BV) festgelegt.

- BV_1: Zu Beginn eines Fahrversuchs erhält der Proband eine Vorstellung aller Bewertungskriterien. Er wird explizit dazu aufgefordert, Unklarheiten anzusprechen.
- BV_2: Er wird explizit aufgefordert, sich besonders auf die Lenkung zu konzentrieren und alle anderen Aspekte nicht weiter zu beachten.
- BV_3: Er wird explizit aufgefordert, über den Versuch hinweg ein möglichst konstantes Fahrverhalten aufzuweisen.
- BV_4: Vor der entsprechenden Fahrsituation wird dem Probanden das zu beurteilende Kriterium genannt und erneut erklärt (Schalz et al., 2002). So ist es ihm möglich, in der Situation konkret darauf zu achten.
- BV_5: Der Proband gibt seine Beurteilung anhand der entwickelten zweistufigen Skala während (Wolf, 2009) oder unmittelbar nach (Rothhämel et al., 2011b) dem Durchfahren der Situation ab und begründet diese.
- BV_6: Der Versuchsleiter wiederholt die Bewertung, um beispielsweise bei invertierten Items Missverständnissen bezüglich der Richtung der Skala vorzubeugen.

Versuchsdesign

Für alle Studien, bei denen verschiedene Fahrzeuge oder Variationen des Lenkgefühls verglichen werden sollen, ist unbedingt ein Versuchsdesign (VD) mit Messwiederholung zu wählen. Dies bedeutet, dass jeder Proband alle zu testenden Alternativen bewertet und somit intra-individuelle Vergleiche möglich sind (Sedlmeier & Renkewitz, 2011). Nur so können inter-individuelle Unterschiede als Störfaktoren eliminiert werden.

- VD_1: Versuchsdesign mit Messwiederholung

Stichprobe

Wie bereits zuvor hervorgehoben, ist das vorliegende Versuchskonzept auf die Bewertung durch Normalfahrer ausgelegt. Die Zusammenstellung der herangezogenen Stichprobe (SP) beeinflusst in einem großen Maß die Qualität der erzielten Ergebnisse. Deshalb werden folgende Anforderungen abgeleitet:

- SP_1: Stichprobengröße mathematisch bestimmt

Einerseits soll die benötigte Stichprobengröße mit dem Programm G*Power (Faul, Erdfelder, Buchner & Lang, 2009; Faul, Erdfelder, Lang & Buchner, 2007) rechnerisch bestimmt werden. Die Berechnung basiert auf der Eingabe verschiedener, für die geplanten Berechnungen relevanter Kennwerte, beispielsweise α -Fehler-Niveau, Power, vermutete Effektstärken, vermutete Korrelationskoeffizienten etc. So kann sowohl eine niedrige Power aufgrund einer zu kleinen Stichprobe als auch eine geringe Versuchsökonomie aufgrund einer zu großen Stichprobe ausgeschlossen werden.

- SP_2: Teilnahmevoraussetzungen: 3 Jahre Führerscheinbesitz, jährliche Fahrleistung von 30.000 km, haupt- oder nebenberuflicher Lkw-Fahrer, Erfahrung mit Lkw-Typ, ausreichende Sprachkenntnisse

Ferner sollten geeignete Teilnahmevoraussetzungen festgelegt werden. Es wird empfohlen, dass Fahrer seit mindestens drei Jahren im Besitz des Lkw-Führerscheins sind und eine jährliche Fahrleistung von 30.000 km aufweisen. Weiterhin sollten sie zum Versuchszeitpunkt einer haupt- oder nebenberuflichen Tätigkeit als Berufskraftfahrer nachgehen. Zusätzlich müssen die Fahrer Erfahrung mit dem im Versuch genutzten Lkw-Typ vorweisen können, also Sattelzug, Gliederzug etc. So soll sichergestellt werden, dass sich die Fahrer auf die Bewertung konzentrieren können. Weiterhin sollte auch auf ausreichende Sprachkenntnisse geachtet werden, damit die Probanden die Kriterien und Erklärungen verstehen können.

- SP_3: Demographische Variablen kontrolliert

Außerdem sollte auch darauf geachtet werden, dass die Fahrer das gesamte Altersspektrum abbilden. Weiterhin ist die Geschlechterverteilung zu beachten. Bei Berufskraftfahrer-Studien mit $N > 1000$ bzw. $N > 2000$ Befragten wird ein Frauenanteil von knapp 5 % (Baier, 2012) bzw. 2 % (ZF Friedrichshafen AG, 2014) berichtet. Da es sich jeweils um anfallende Stichproben handelte, ist von einer Übertragbarkeit auf die tatsächliche Population auszugehen.

3.1.3. Entwurf des Versuchskonzepts

Fragebogen

Der erste Entwurf des Fragebogens zur subjektiven Bewertung des Lenkgefühls schwerer Nutzfahrzeuge durch Normalfahrer besteht aus 35 Bewertungskriterien mit den dazugehörigen Erklärungen (Tabelle 14). Weiterhin werden die Fahrsituationen, in denen die einzelnen Bewertungskriterien erlebt werden können, spezifiziert.

Tabelle 14. Erster Entwurf des Fragebogens.

Nr.	Bewertungskriterium*	Erklärung	Spezifische Fahrsituation
1	Agil	Das Fahrzeug reagiert sehr willig auf Lenkbewegungen.	Mittlere Lenkwinkel in beide Richtungen bei mittleren bis hohen Geschwindigkeiten
2	Anlenkmoment	Zum Anlenken ist ein geringer Kraftaufwand notwendig.	Kleine Lenkwinkel bei hohen Geschwindigkeiten
3	Beherrschbar	/	Engstelle
4	Direkt	Bereits ein kleines Einlenken bewirkt eine große Fahrzeugreaktion.	Alle Lenkwinkelbereiche in allen Geschwindigkeitsbereichen
5	Empfindlich (i)	Fahrbahneinflüsse sind am Lenkrad spürbar.	Fahrbahnoberfläche (Fahrbahnmarkierung, Querfuge)
6	Flattert (i)	Das Lenkrad zappelt.	Geradeausfahrt bei hohen Geschwindigkeiten
7	Genau	Das Fahrzeug reagiert genau auf Ihre Lenkbewegungen.	Engstelle
8	Geradeauslauf	Das Fahrzeug fährt im Geradeauslauf wie auf Schienen.	Geradeausfahrt bei mittleren und hohen Geschwindigkeiten
9	Geschmeidig	/	Global
10	Gleichmäßig	Der Kraftaufwand ist konstant über alle Geschwindigkeiten und Lenkwinkelbereiche hinweg. Sie können an keiner Stelle des Einlenkens einen erhöhten Kraftbedarf feststellen.	Große Lenkwinkel bei niedrigen Geschwindigkeiten
11	Hakt (i)	An manchen Stellen stockt die Lenkung, sodass Sie plötzlich mehr Kraft aufwenden müssen.	Große Lenkwinkel bei niedrigen Geschwindigkeiten
12	Haltemoment	Fürs Halten der Lenkung in Kurven ist ein geringer Kraftaufwand nötig.	Konstanter Lenkwinkel
13	Hat Spiel (i)	Es gibt einen toten Bereich um die Mitte.	Geradeausfahrt bei hohen Geschwindigkeiten
14	Klebrig (i)	/	Global
15	Knarzt (i)	Sie können ein Lenkgeräusch hören und evtl. fühlen.	Große Lenkwinkel bei niedrigen Geschwindigkeiten
16	Komfortabel	/	Global
17	Korrekturbedarf (i)	Sie müssen häufig nachkorrigieren.	Kleiner Lenkwinkel bei mittleren Geschwindigkeiten
18	Leichtgängig	Um zu lenken ist nur ein geringer Kraftaufwand notwendig.	Große Lenkwinkel bei niedrigen Geschwindigkeiten
19	Lenkaktivität (i)	Sie müssen beim Lenken viel kurbeln.	Große Lenkwinkel in beide Richtungen bei niedrigen Geschwindigkeiten

20	Lenkungsstößigkeit (i)	Die Lenkung leitet Störungen der Straße weiter (z. B. Kanaldeckel, Querfuge).	Schlechte Fahrbahnoberfläche (Kanaldeckel, ausgeprägte Teerfugen)
21	Mittenzentrierung	Die Geradeausstellung des Lenkrads ist fühlbar.	Geradeausfahrt bei mittleren und hohen Geschwindigkeiten
22	Nervös (i)	Das Fahrzeug reagiert hektisch und unerwartet.	Geradeausfahrt bei hohen Geschwindigkeiten
23	Rückstellverhalten	Die Lenkung stellt sich ausreichend selbsttätig in die Geradeausstellung zurück.	Nach großen Lenkwinkeln bei niedrigen Geschwindigkeiten
24	Ruhig	/	Global
25	Schlackert (i)	Sie haben den Eindruck, dass die Räder wackeln.	Geradeausfahrt bei hohen Geschwindigkeiten
26	Schwammig (i)	/	Global
27	Schwimmt (i)	Die Fahrzeugreaktion ist unscharf (ähnlich zu glatter Fahrbahn).	Kleine Lenkwinkel in beide Richtungen bei mittleren und hohen Geschwindigkeiten
28	Sicher	/	Global
29	Spurgetreu	Es ist wenig Aufwand nötig, um das Fahrzeug in der Spur zu halten.	Kleiner Lenkwinkel bei mittleren Geschwindigkeiten
30	Stabil	/	Global
31	Straff	/	Mittlere Lenkwinkel in beide Richtungen bei mittleren bis hohen Geschwindigkeiten
32	Synthetisch (i)	Die Lenkung liefert keine Rückmeldung des Fahrzustandes (z. B. Radstellung).	Fahrbahnoberfläche (Fahrbahnmarkierung, Querfuge)
33	Verzögert (i)	Das Fahrzeug reagiert spät auf Ihre Lenkbewegungen.	Mittlere Lenkwinkel in beide Richtungen bei mittleren bis hohen Geschwindigkeiten
34	Vibriert (i)	Es sind schnelle Schwingungen am Lenkrad spürbar.	Geradeausfahrt bei hohen Geschwindigkeiten
35	Zäh (i)	/	Mittlere Lenkwinkel in beide Richtungen bei mittleren bis hohen Geschwindigkeiten

* (i): invertiert

Die Bewertungskriterien werden um eine Bewertungsskala ergänzt, anhand welcher eine subjektive Einschätzung vorgenommen wird. Die Skalen zur absoluten und relativen Bewertung sind in Abbildung 10 zu finden.

Absolute Bewertung

Zeigt das Fahrzeug / die Variante diese Eigenschaft?										
1	2	3	4	5	6	7	8	9	10	
gar nicht										völlig
Gefällt Ihnen, wie stark diese Eigenschaft ausgeprägt ist?										
1	2	3	4	5	6	7	8	9	10	
gar nicht										völlig

Relative Bewertung

Wie stark zeigt das Fahrzeug / die Variante diese Eigenschaft im Vergleich zum vorher bewerteten Fahrzeug / zur vorher bewerteten Variante / zur Referenz?										
-5	-4	-3	-2	-1	0	1	2	3	4	5
sehr viel weniger										sehr viel mehr
Wie sehr gefällt Ihnen, wie stark diese Eigenschaft ausgeprägt ist, im Vergleich zum vorher bewerteten Fahrzeug / zur vorher bewerteten Variante / zur Referenz?										
-5	-4	-3	-2	-1	0	1	2	3	4	5
sehr viel weniger										sehr viel mehr

Abbildung 10. Skalen zur absoluten und relativen Niveau- und Gefallen-Bewertung.

Versuchsaufbau

Die Anforderungen zum Versuchsaufbau sind als Checkliste zu verstehen, welche bei der Konzeption einer Studie herangezogen werden kann (Tabelle 15). Die vorliegenden Anforderungen gliedern sich in die Bereiche Fahrsituationen, Bewertungsvorgehen, Versuchsdesign und Stichprobe.

Tabelle 15. Checkliste zum Versuchsaufbau.

Bereich	Anforderungen
Fahrsituationen (FS)	<input type="checkbox"/> FS_1: Gesamter Fahrbetrieb abgebildet <input type="checkbox"/> FS_2: Fahrsituationen entsprechen fahrerischem Können der Normalfahrer <input type="checkbox"/> FS_3: Fahrzeit möglichst kurz <input type="checkbox"/> FS_4: Zu Bewertungskriterien zugehörige Fahrsituationen spezifiziert
Bewertungsvorgehen (BV)	<input type="checkbox"/> BV_1: Vorstellung aller Bewertungskriterien vor der Fahrt <input type="checkbox"/> BV_2: Aufforderung an Proband, Konzentration auf Lenkung zu richten <input type="checkbox"/> BV_3: Aufforderung an Proband, konstantes Fahrverhalten aufzuweisen <input type="checkbox"/> BV_4: Erklären des Kriteriums vor korrespondierender Fahrsituation <input type="checkbox"/> BV_5: Begründete Beurteilung anhand Skala während oder nach Fahrsituation <input type="checkbox"/> BV_6: Versuchsleiter wiederholt Bewertung
Versuchsdesign (VD)	<input type="checkbox"/> VD_1: Versuchsdesign mit Messwiederholung
Stichprobe (SP)	<input type="checkbox"/> SP_1: Stichprobengröße mathematisch bestimmt <input type="checkbox"/> SP_2: Teilnahmevoraussetzungen: 3 Jahre Führerscheinbesitz, jährliche Fahrleistung von 30.000 km, haupt- oder nebenberuflicher Lkw-Fahrer, Erfahrung mit Lkw-Typ, ausreichende Sprachkenntnisse <input type="checkbox"/> SP_3: Demographische Variablen kontrolliert

3.2. Erste Evaluation des Versuchskonzepts: Bewertung des Lenkgefühls verschiedener Fahrzeuge (Studie I)

3.2.1. Ziele und Hypothesen

Das entwickelte Versuchskonzept wurde anschließend in einem Fahrversuch angewendet. Ziel des Versuchs war zum einen die Evaluation des Fragebogens. Hierzu wurde vor der Berechnung der Hypothesentests eine explorative Faktorenanalyse durchgeführt. So wurden die psychometrischen Gütekriterien des Fragebogens bestimmt und die Gültigkeit der gewonnenen Ergebnisse überprüft. Zum anderen sollte untersucht werden, inwiefern sich die Bewertungen der Normalfahrer, die die Zielgruppe dieses Versuchskonzepts darstellen, von denen von Experten unterscheiden.

Die subjektiven Bewertungen des Lenkgefühls dienten in diesem Versuch lediglich zur Evaluation des Versuchskonzepts, nicht aber zur Analyse des Lenkgefühls selbst. Deshalb wurde eine möglichst leicht realisierbare Operationalisierung gewählt (Kapitel 2.3.2). So wurde das Lenkgefühl mithilfe von zwei Sattelzügen variiert, welche sich objektiv hinsichtlich mehrerer Aspekte der Lenkungsauslegung unterscheiden. Die Evaluation fand im öffentlichen Straßenverkehr statt.

Wie Bewertungen im Allgemeinen könnte auch die Bewertung des Lenkgefühls durch verschiedene Faktoren beeinflusst werden, wie bspw. Geschlecht und Alter (Ackert, 2008). Hier ist z. B. die unterschiedliche Optik der verwendeten Fahrzeuge oder eine Favoritenbildung im Laufe des Versuchs zu nennen (Gies & Marusic, 2000).

Außerdem können die Fahrer bereits im Vorfeld eine bestimmte Meinung gegenüber den Fahrzeugen oder Herstellern haben, welche durch die Medien (Schalz et al., 2002) oder eigene Erfahrungen geprägt sind. Somit können die Bewertungen des Lenkgefühls verschiedener Fahrzeuge durch die Einstellung der Fahrer beeinflusst werden. Mit dem vorliegenden Versuchskonzept soll es dennoch möglich sein, von Probanden zuverlässige Erkenntnisse zum Lenkgefühl zu sammeln. Dies soll auch dann zutreffen, wenn sie bereits Erfahrung mit den Versuchsfahrzeugen haben. Hieraus leitet sich Hypothese 1 ab:

H1: Die Niveau- und Gefallen-Bewertungen der Fahrzeuge sind unabhängig davon, welches Fahrzeug die Probanden üblicherweise fahren.

Anschließend sollte überprüft werden, ob Normalfahrer und Experten das Lenkgefühl der beiden Fahrzeuge gleichermaßen bewerten. Experten versetzen sich im Auslegungsprozess stets in die Kunden hinein und möchten ihre Wünsche abbilden (Wolf, 2009). Da das Expertenurteil dem der Normalfahrer entsprechen sollte, lautet Hypothese 2:

H2: Die Bewertungen der Normalfahrer und Experten sind nicht signifikant verschieden.

Nichtsdestotrotz wurde im Stand der Forschung (Kapitel 2.3.1) vorgestellt, dass Experten aufgrund ihrer Fähigkeiten und Fertigkeiten differenziertere und konsistentere Beurteilungen als Normalfahrer abgeben können. Hieraus folgt für Hypothese 3:

H3: Die Bewertungen der Experten sind konsistenter als die der Normalfahrer.

Zur Analyse der Differenziertheit der Beurteilungen wurde die Effektstärke genutzt, um Aussagen zur Größe der wahrgenommenen Unterschiede zu machen. Als Maß der Streuung von Beurteilungen wurde die Standardabweichung herangezogen.

H3a: Die wahrgenommenen Unterschiede fallen bei den Experten größer aus als bei den Normalfahrern.

H3b: Die Bewertungen der Experten weisen kleinere Standardabweichungen auf als die der Normalfahrer.

Die Vorbereitung, Durchführung und Auswertung der Studie wurde von Schmidt (2014) im Rahmen ihrer Bachelorthesis unterstützt.

3.2.2. Methodik

Versuchsdesign

Um oben genannte Hypothesen zu überprüfen wurde ein zweifaktorielles Versuchsdesign genutzt (Tabelle 16). Der Faktor *Fahrzeug* hatte die zwei Stufen Fahrzeug 1 (FZG1) und Fahrzeug 2 (FZG2). Das Versuchsdesign mit Messwiederholung (within-subject) wurde dadurch realisiert, dass alle Probanden beide Fahrzeuge erprobten. Der Faktor *Sub-Stichprobe* hatte vier Stufen (between-subject). Die Sub-Stichproben 1-3 wurden mit Normalfahrern besetzt. Gemäß Hypothese 1 waren Probanden zu rekrutieren, die Erfahrung mit FZG1 bzw. FZG2 hatten. Um die Anzahl der in Frage kommenden Probanden zu erhöhen, war auch Erfahrung mit einem Fahrzeug des entsprechenden Herstellers akzeptabel. Sub-Stichprobe 1 bestand aus Fahrern, die Fahrerfahrung mit FZG1 oder dessen Hersteller vorweisen konnten (F_HRST1). Für Sub-Stichprobe 2 galt dies in Bezug auf FZG2 (F_HRST2). Trotz einer entsprechenden Vorauswahl stellte sich bei einigen Fahrern heraus, dass sie Erfahrung mit beiden

Fahrzeugen bzw. ihren Herstellern oder mit keinem von beiden hatten. Diese Fahrer wurden nicht ausgeschlossen, sondern als Sub-Stichprobe 3 aufgenommen (F_HRST3). Sub-Stichprobe 4 war mit Experten besetzt (EXP).

Die abhängigen Variablen waren die Niveau- und Gefallen-Bewertungen der Probanden zu den einzelnen Bewertungskriterien sowie eine Gesamt-Gefallenbewertung. Zusätzlich wurden die verbalen Äußerungen der Probanden während der Fahrt sowie eine Versuchsleitereinschätzung zum Verständnis der Probanden notiert. Außerdem wurde außerdem etwas mehr als die Hälfte der Normalfahrer ($n = 23$) zum Verständnis des Fragebogens nachbefragt.

Tabelle 16. Versuchsdesign unter Angabe der Stichprobengrößen.

Faktor 2: Sub-Stichprobe	Faktor 1: Fahrzeug	
	FZG1	FZG2
Fahrer von Hersteller 1 (F_HRST1)	$n = 27$	$n = 27$
Fahrer von Hersteller 2 (F_HRST2)	$n = 6$	$n = 6$
Fahrer von Hersteller 1 & 2 oder Fahrer anderer Hersteller (F_HRST3)	$n = 8$	$n = 8$
Experten (EXP)	$n = 10$	$n = 10$

Stichprobe

Die Stichprobe der Normalfahrer wurde gemäß der Anforderungen des Versuchskonzepts zusammengestellt: Die benötigte Stichprobengröße wurde mit G*Power 3.1.9.2 (Faul, 2007; Faul, 2009) festgesetzt. Zur Überprüfung von Hypothese 1 sollten t -Tests für unabhängige Stichproben berechnet werden. Um mit einer angenommenen mittleren Effektstärke von $d = .5$ und einer Power von $(1-\beta) = .8$ einen tatsächlich vorhandenen Effekt aufzudecken, wurden $n = 64$ Probanden pro Sub-Stichprobe benötigt. Für alle weiteren Hypothesen wurden die Daten mehrerer Bewertungskriterien zusammengelegt, sodass hier die zuvor bestimmte Stichprobengröße ausreicht. Da es sich um eine erste Evaluation des Versuchskonzepts handelte, widersprach die Erhebung einer derart großen Stichprobe jedoch der Versuchsökonomie. Somit sollten als Kompromiss $N = 40$ auswertbare Datensätze vorliegen. Die Teilnehmer wurden aus einem bestehenden Probandenpool rekrutiert und erhielten eine monetäre Aufwandsentschädigung.

Insgesamt wurden $N = 45$ externe Fahrer rekrutiert, wobei vier Probanden aufgrund zu geringer Deutsch-Kenntnisse vor den Analysen ausgeschlossen werden mussten. Von den $N = 41$ auswertbaren Fahrern waren zwei Probanden weiblich (4,9 %). Angaben zum Alter sowie der Fahrerfahrung finden sich in Tabelle 17. Ein Anteil von 65,9 % der Fahrer war im Verteilerverkehr tätig, 31,7 % im Fernverkehr. 70,7 % waren hauptberuflich Kraftfahrer, 29,3 % nebenberuflich. Fünf Fahrer gaben an, im regelmäßigen Fahrbetrieb keinen Sattelzug zu fahren.

Tabelle 17. Alter und Fahrpraxis der Normalfahrer.

Variable	<i>Min.</i>	<i>Max.</i>	<i>M</i>	<i>SD</i>
Alter [Jahre]	24	62	43,3	9,2
Fahrleistung [km]	5.000	180.000	67.695	41.955
Fahrerfahrung [Jahre]	3	43	20,8	10,0

$n = 41$

Anschließend wurde das Fahrverhalten hinsichtlich der üblicherweise gefahrenen Fahrzeuge analysiert. Fahrer, die mindestens 30 % der Zeit ein Fahrzeug eines bestimmten Herstellers fuhren, wurden als Fahrer dieses Herstellers (F_HRST) bezeichnet. Für die drei Sub-Stichproben ergaben sich folgende Anteile:

- F_HRST1: $n = 27$ (65,9 %)
- F_HRST2: $n = 6$ (14,6 %)
- F_HRST3: $n = 8$ (19,5 %)

Somit wurde keine Gleichverteilung der Stichprobengrößen von F_HRST1 und F_HRST2 erreicht.

Wie oben angesprochen nahmen auch $N = 10$ Experten teil. Für sie ist nach Bubb (2003) eine derart kleine Stichprobengröße ausreichend, da aufgrund ihrer Expertise von zuverlässigeren Beurteilungen auszugehen ist. Die Experten hatten mehrjährige Erfahrung in der Lenkungsauslegung und subjektiven Lenkungsbeurteilung. Sechs Experten waren Ingenieure, vier Werkstattmitarbeiter. Angaben zum Alter sowie der Fahrerfahrung finden sich in Tabelle 18. Alle Experten waren männlich.

Tabelle 18. Alter und Fahrpraxis der Experten.

Variable	<i>Min.</i>	<i>Max.</i>	<i>M</i>	<i>SD</i>
Alter [Jahre]	30	59	43,3	11,0
Fahrleistung [km]	4.000	100.000	25.550	31.846
Fahrerfahrung [Jahre]	4	38	17,6	11,1

$n = 10$

Variation des Lenkgefühls

Die Variation des Lenkgefühls erfolgte mithilfe verschiedener Fahrzeuge. Es wurden zwei aktuelle Fabrikate genutzt, von welchen bekannt ist, dass sie sich hinsichtlich ihres Lenkgefühls deutlich unterscheiden. So sollte es den Probanden möglichst einfach gemacht werden, Unterschiede in der Lenkung zu erkennen. Zur Auswahl geeigneter Fahrzeuge wurden die Ergebnisse einer internen Untersuchung gesichtet, in welcher mithilfe verschiedener standardisierter Fahrmanöver und Messverfahren die Fahrdynamik von Sattelzugmaschinen

verglichen wurde. Ferner wurden Experten-Einschätzungen aus einem Prüfbericht (Koschinsky, 2014) sowie Presse-Berichte (Becker, 2011; Grünig, 2013; Zeitzen, 2012, 2013) herangezogen.

Auf Basis dieser Quellen wurden Bewertungskriterien identifiziert, bei denen die Auslegung von FZG1 deutlich besser ist als die von FZG2 (*leichtgängig (Rangieren), Lenkungsstößigkeit*). Andererseits wurden Bewertungskriterien identifiziert, bei denen die Auslegung von FZG2 deutlich besser ist als die von FZG1 (*agil, direkt, hat Spiel*). Für einige weitere Kriterien waren ebenfalls positivere Bewertungen für FZG1 (*empfindlich, Stabilität*) bzw. FZG2 (*leichtgängig (Fahrt), schwammig, verzögert*) zu erwarten. Bei der Gesamt-Gefallenbewertung wurden für FZG2 bessere Werte erwartet als für FZG1.

Um die Randbedingungen möglichst konstant zu halten, wurden beide Zugmaschinen mit identischen, voll beladenen Euro-Aufliegern gekoppelt. Radstand und Bereifung beider Fahrzeuge waren identisch. Sie wiesen ferner einen vergleichbaren Gesamt-Kilometerstand von ca. 100.000 km auf. Es lagen jedoch, wie bei der Nutzung verschiedener Fahrzeuge unumgänglich, einige Unterschiede vor. Hier sind beispielsweise die Fahrerhaushöhe sowie die Vorder- und Hinterachse zu nennen. Ferner war auch die Gesamtlenkübersetzung verschieden.

Versuchsumgebung und Fahrsituationen

Die Studie sollte im Straßenverkehr durchgeführt werden.

Die Versuchsstrecke sollte maximal 1 h Fahrt in Anspruch nehmen, um die Konzentrationsfähigkeit der Probanden über die gesamte Dauer aufrecht zu erhalten. Die Teststrecke wurde so ausgewählt, dass sie mit dem Lkw möglichst problemlos durchfahrbar war. Ferner sollte sie wenig stauanfällig sein, sodass Umgebungsfaktoren wenig Einfluss auf Fahrgeschwindigkeit und Fahrverhalten der Probanden hatten. Außerdem war zu beachten, dass alle zur Beurteilung nötigen Fahrsituationen vorhanden waren. Neben Rangier-, Stadt-, Land- und Autobahnabschnitten musste auch eine ausreichende Anzahl an kurvigen bzw. geraden Abschnitten, Kreisverkehren, Abbiegesituationen und Strecken mit schlechter Fahrbahnoberfläche vorhanden sein.

Die ausgewählte Strecke war 30 km lang und führte nach einem kurzen Abschnitt in der Stadt über einen Autobahnabschnitt, gefolgt von sich abwechselnden Stadt- und Landstraßenfahrten. Die Rangiersituation wurde auf dem Werksgelände absolviert. Die Bewertungskriterien wurden über die Versuchsstrecke verteilt geeigneten Fahrsituationen zugeordnet. Dieses Vorgehen ist in Abbildung 11 anhand eines Abschnitts der Versuchsstrecke dargestellt. Einige Kriterien wurden zwei Mal bewertet. Insgesamt wurden für eine Runde ca. 55 min benötigt.

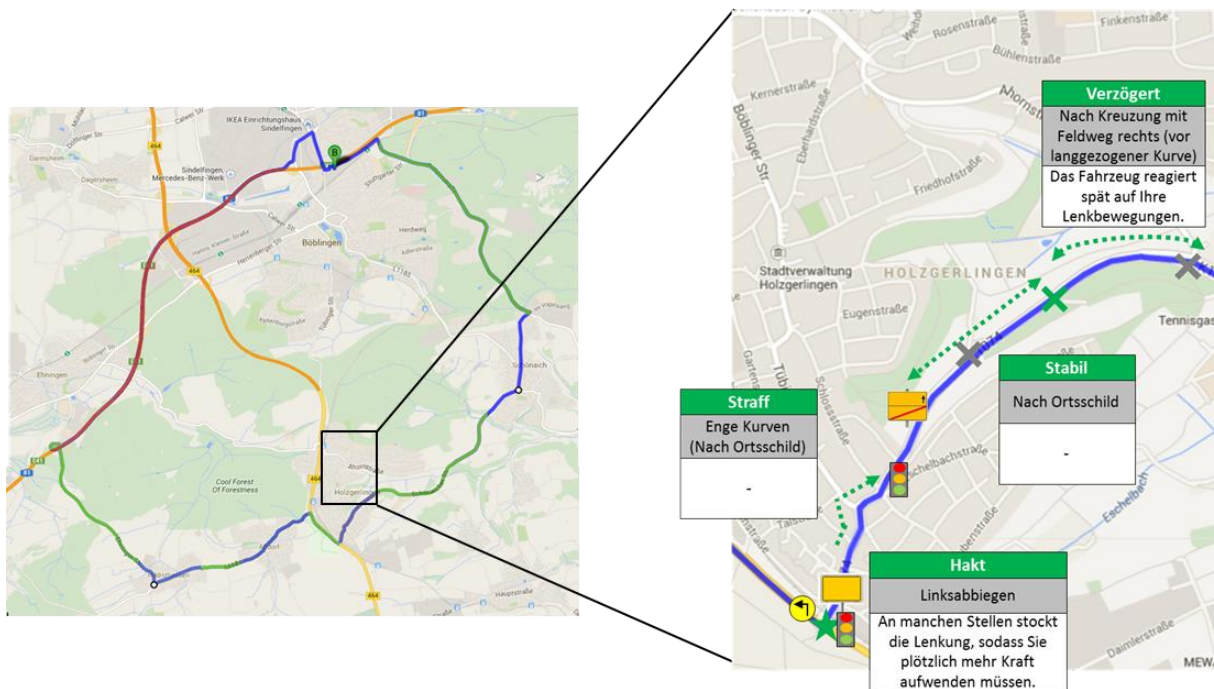


Abbildung 11. Versuchsstrecke mit beispielhaftem Streckenabschnitt zur Verdeutlichung des Bewertungsvorgehens (Kartenmaterial: Google Maps, 2014).

Art der Bewertung, Bewertungsskala und Bewertungskriterien

Wie in Kapitel 2.3.2 bzw. Tabelle 4 gezeigt, war bei der Kombination aus der gewählten Variation des Lenkgefühls und der Versuchsumgebung nur eine absolute Bewertung möglich. Dementsprechend wurde zur Bewertung die 10-stufige Absolut-Skala genutzt (Kapitel 3.1.3).

Versuchsablauf

Der Versuch dauerte insgesamt ca. 2,5 h pro Proband. Zuerst fand eine Vorbesprechung statt, um den Teilnehmern einen Einblick in die bevorstehende Beurteilungsaufgabe zu geben. Dazu wurde der Kriterienkatalog inklusive der Erklärungen erläutert. Die Probanden sollten eventuell auftretende Unklarheiten ansprechen, da sich der Fragebogen noch in der Entwicklung befand. Im Anschluss wurde die Bewertung anhand der Aspekte Niveau und Gefallen erläutert und die Beurteilung mit der 10-stufigen Skala an einem Beispiel geübt. Darauf folgte die Erhebung der demographischen Daten sowie Fahrgewohnheiten.

Anschließend wurde die oben vorgestellte Strecke zwei Mal absolviert, wobei die Reihenfolge der Versuchsfahrzeuge randomisiert war. Die Probanden wurden während der beiden Fahrten von einem Versuchsleiter befragt. Auch jetzt sollten sie auf eventuelle Schwierigkeiten hinweisen. An den zuvor festgelegten Stellen wurden die einzelnen Kriterien gemäß dem Bewertungsvorgehen beurteilt. Die Bewertungen waren absolut, jedoch nannte der Versuchsleiter bei der zweiten Fahrt auf Nachfrage die Bewertung des ersten Fahrzeugs. Qualitative Anmerkungen des Probanden sowie aufgetretene Verständnisprobleme wurden vom Versuchsleiter notiert. Am Ende jeder Fahrt wurde die Gesamtbewertung abgegeben und begründet. Nach der zweiten Fahrt wurden mit den Teilnehmern die von ihnen wahrgenommenen Unterschiede zwischen den

Lenkungen der beiden Versuchsfahrzeuge herausgearbeitet. Ein Teil der Normalfahrer gab abschließend an, welche der Kriterien sie besonders gut verstanden haben.

Auswertungsvorgehen

Im ersten Schritt sollten nicht verstandene Kriterien identifiziert werden. Hierzu wurden einerseits die von den Versuchsleitern während der Fahrt wahrgenommenen oder von den Probanden geäußerten Verständnisschwierigkeiten analysiert. Andererseits wurde die Nachbefragung genutzt. Es wurden fünf unverständliche Kriterien entdeckt, beispielsweise *klebrig* und *straff*. Da diese nicht elementar für die Beschreibung des Lenkgefühls waren, wurden sie eliminiert. Im Anschluss wurden die Bewertungen der doppelt erhobenen Kriterien gemittelt.

Die statistische Auswertung erfolgte mit IBM SPSS Statistics 21. Die Niveau-Bewertungen der invertierten Kriterien wurden umgepolt. Somit standen bei allen Kriterien höhere Niveau-Bewertungen für eine positivere Ausprägung des Lenkgefühls. Die Daten der Niveau- und Gefallen-Bewertungen sowie der Gesamt-Gefallenbewertung wurden dann einer deskriptiven Analyse unterzogen. Da sich Fahrer in der Nutzung der Skala stark unterscheiden (Krüger & Neukum, 2001), sollten die Daten transformiert werden. Transformationen im Allgemeinen werden von Bush, Hess und Wolford (1993) empfohlen. Sie werden besonders oft angewendet, wenn die vorliegenden Daten die Voraussetzungen für die gewünschte Analyse nicht erfüllen. Außerdem eliminieren sie bei Versuchsdesigns mit Messwiederholung Unterschiede im Antwortverhalten zwischen den Personen. Insbesondere mit einer z -Transformation wurden über alle vorliegenden Verteilungen hinweg Verbesserungen erzielt (Bush et al., 1993). Die z -Transformation wurde bereits in einigen Studien zum Lenkgefühl angewendet (Chen & Crolla, 1998; Gil Gómez et al., 2015; Kim & Yoon, 2015; Zschocke, 2009). Auf die vorliegenden Daten wird sie gemäß folgender Formel angewendet (Bortz & Schuster, 2010):

$$z = \frac{x - \bar{x}}{s}$$

Der transformierte z -Wert wird aus dem ursprünglichen Wert x eines Probanden, dem Mittelwert \bar{x} über die Bewertungen aller Probanden innerhalb dieses Kriteriums und der Standardabweichung s berechnet.

Entsprechend der Voraussetzungen wurden die Daten anschließend mithilfe von Kolmogorov-Smirnov-Tests auf Normalverteilung geprüft. Nach Bortz und Schuster (2010) ist eine Verletzung der Normalverteilung bei $N \geq 30$ gemäß des zentralen Grenzwertsatzes vernachlässigbar. Somit wurden lediglich die Daten aus Stichproben mit $N < 30$ überprüft. Waren die Voraussetzungen für parametrische Verfahren nicht erfüllt, wurden die entsprechenden nicht-parametrischen Verfahren angewendet.

Das Signifikanzniveau betrug $\alpha = .05$. Signifikanzwerte unter .05 werden mit * gekennzeichnet, Werte unter .01 mit ** und Werte unter .001 mit ***. Es ist gängige Auswertungspraxis, den p -Wert bei gerichteten Hypothesen zu halbieren. Einige Studien kamen aber zu dem Ergebnis, dass das Durchführen einseitiger Tests aufgrund nicht erfüllter Voraussetzungen häufig zu Unrecht erfolgt (Lombardi & Hurlbert, 2009; Ruxton & Neuhäuser, 2010). Deshalb wurden, wie bei Field (2013) empfohlen, zweiseitige Tests durchgeführt. Zur Vermeidung der α -Fehler-Kumulierung wurde jeweils eine Benjamini-Hochberg-Korrektur für die Hypothesentests bzw. die explorative

Analyse durchgeführt (Benjamini & Hochberg, 1995). Dieses Verfahren kontrolliert die false discovery rate (FDR). Auch wenn die Methode ursprünglich für unabhängige Tests entwickelt wurde, wurde sie bereits im Hinblick auf abhängige Daten untersucht (Kim & van de Wiel, 2008; Yekutieli & Benjamini, 2001).

Für paarweise Vergleiche wurden Effektstärken berechnet. Es gelten gemäß Cohen (1988) für r die Grenzen .1 für einen kleinen, .3 für einen mittleren und .5 für einen großen Effekt.

Die psychometrische Untersuchung des Fragebogens erfolgte mithilfe einer explorativen Faktoren-, Item- und Reliabilitätsanalyse.

3.2.3. Ergebnisse

Hypothese 1: Unabhängigkeit der Bewertungen vom üblicherweise gefahrenen Fahrzeug

H1: Die Niveau- und Gefallen-Bewertungen der Fahrzeuge sind unabhängig davon, welches Fahrzeug die Probanden üblicherweise fahren.

In Kapitel 3.2.2 wurden Ergebnisse aus einem Prüfbericht, einer Experten-Evaluation sowie Presseberichten zur Bewertung des Lenkgefühls der beiden Fahrzeuge FZG1 und FZG2 zusammengefasst. Aus diesen lässt sich ableiten, dass für bestimmte Bewertungskriterien bei FZG1 bessere Niveau- und Gefallen-Bewertungen zu erwarten waren als bei FZG2. Andererseits gab es Bewertungskriterien, für die bei FZG2 bessere Niveau- und Gefallen-Bewertungen zu erwarten waren als bei FZG1. Bezüglich der Gesamt-Gefallenbewertung waren höhere Werte für FZG2 zu erwarten.

Aufgrund der Stichprobengrößen von $n = 27$ bei F_HRST1 und $n = 6$ bei F_HRST2 war ein unmittelbarer Vergleich der beiden Sub-Stichproben nicht sinnvoll. Vielmehr wurden die Bewertungen der FZG1 und FZG2 innerhalb der Sub-Stichprobe F_HRST1 verglichen. Diese Analysen wurden für die Bewertungen innerhalb der Sub-Stichprobe F_HRST2 wiederholt. Es wurden nicht-parametrische Wilcoxon-Vorzeichen-Rang-Tests berechnet. Die deskriptiven Statistiken sind in Anhang I zu finden.

Zuerst wird auf die Bewertungskriterien eingegangen, bei denen von besseren Bewertungen von FZG1 auszugehen war. Die Sub-Stichprobe F_HRST1 beurteilte die *Lenkungsstößigkeit* von FZG1 besser als die von FZG2, was sich in entsprechend höheren Gefallen-Bewertungen äußerte (Tabelle 19 und Abbildung 12). Den Unterschieden lagen große Effektstärken zu Grunde. Bei der Sub-Stichprobe F_HRST2 lagen keine Unterschiede zwischen den Bewertungen der beiden Versuchsfahrzeuge vor.

Tabelle 19. Ergebnisse der Wilcoxon-Vorzeichen-Rang-Tests zum Vergleich der Bewertungen von *leichtgängig* und *Lenkungsstößigkeit* für die Sub-Stichproben F_HRST1 und F_HRST2 getrennt.

Bewertungskriterium	Sub-Stichprobe	z	p	r
Leichtgängig (Rangieren) – Niveau	F_HRST1	-1,55	.201	.21
	F_HRST2	-0,14	.949	.04
Leichtgängig (Rangieren) – Gefallen	F_HRST1	-2,21	.069	.30
	F_HRST2	0,00	1.000	.00
Lenkungsstößigkeit – Niveau	F_HRST1	-4,06	< .001	.55
	F_HRST2	-1,29	.283	.37
Lenkungsstößigkeit – Gefallen	F_HRST1	-4,00	< .001	.54
	F_HRST2	-0,96	.444	.28

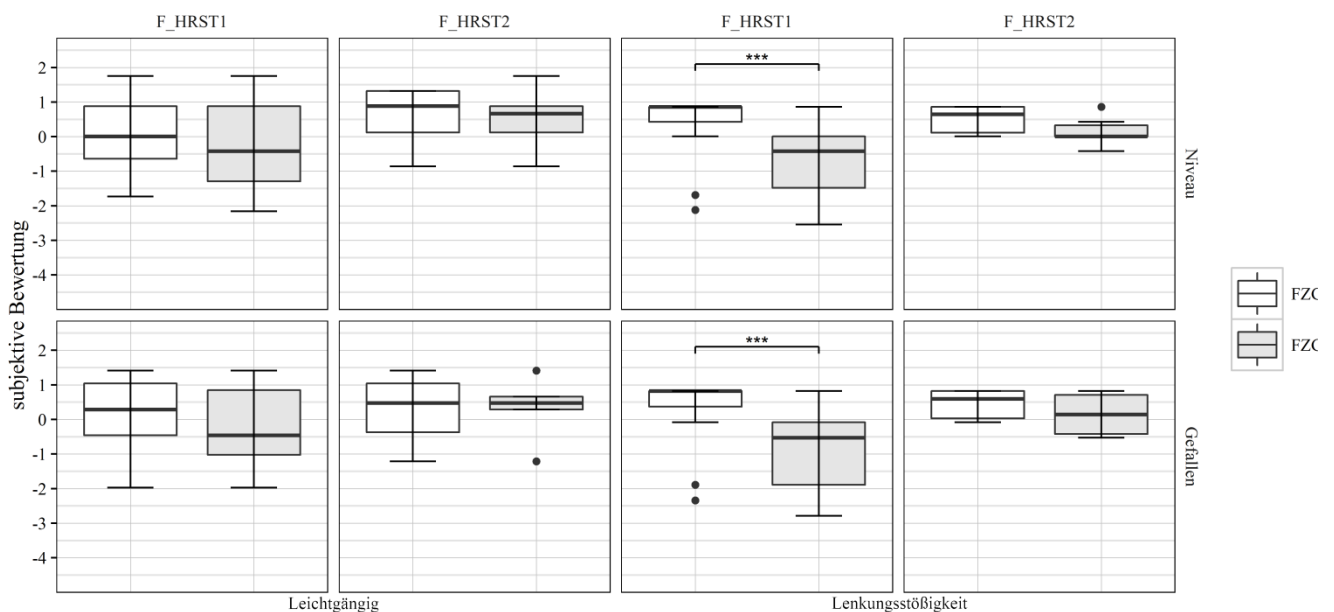


Abbildung 12. Die Mittelwerte der Niveau- und Gefallen-Bewertungen von *leichtgängig* und *Lenkungsstößigkeit* für die Sub-Stichproben F_HRST1 und F_HRST2 getrennt. Die Fehlerbalken repräsentieren die Standardabweichung.

Anschließend wird auf die Bewertungskriterien eingegangen, bei denen von besseren Bewertungen von FZG2 auszugehen war. Bei der Sub-Stichprobe F_HRST1 lagen signifikante Unterschiede zwischen den Gefallen-Bewertungen von FZG1 und FZG2 bei den Kriterien *agil* und *direkt* vor (Tabelle 20 und Abbildung 13). In beiden Fällen wurde, entgegen der Richtung des vermuteten Unterschieds, FZG1 besser bewertet als FZG2. Bei der Sub-Stichprobe F_HRST2 lagen keine Unterschiede zwischen den Bewertungen der beiden Versuchsfahrzeuge vor. Vor der α -Fehler-Korrektur waren jedoch drei der sechs Vergleiche signifikant. Außerdem waren auch die

gefundenen Effektstärken im mittleren bis hohen Bereich. Deshalb wurde eine deskriptive Analyse angeschlossen. Diese zeigte, dass FZG2 in allen sechs Vergleichen deskriptiv höhere Bewertungen erhielt als FZG1.

Abschließend wurde untersucht, ob die Gesamt-Gefallenbewertungen von FZG1 besser sind als die von FZG2. Entgegen der Erwartung bewertete die Sub-Stichprobe F_HRST1 FZG1 hinsichtlich des Gesamt-Gefallens signifikant besser als FZG2. Die Sub-Stichprobe F_HRST2 bewertete die Fahrzeuge nicht signifikant verschieden; die zugrundeliegende Effektstärke war vernachlässigbar niedrig.

Tabelle 20. Ergebnisse der Wilcoxon-Vorzeichen-Rang-Tests zum Vergleich der Bewertungen von *agil*, *direkt*, *hat Spiel* und der Gesamt-Gefallenbewertung für die Sub-Stichproben F_HRST1 und F_HRST2 getrennt.

Bewertungskriterium	Sub-Stichprobe	<i>z</i>	<i>p</i>	<i>r</i>
Agil – Niveau	F_HRST1	-0,92	.453	.13
	F_HRST2	-1,63	.188	.47
Agil – Gefallen	F_HRST1	-2,77	.017	.38
	F_HRST2	-0,82	.506	.24
Direkt – Niveau	F_HRST1	-1,42	.246	.20
	F_HRST2	-2,04	.091	.59
Direkt – Gefallen	F_HRST1	-2,78	.018	.38
	F_HRST2	-2,03	.087	.59
Hat Spiel – Niveau	F_HRST1	-1,18	.325	.16
	F_HRST2	-2,12	.080	.61
Hat Spiel – Gefallen	F_HRST1	-1,39	.246	.19
	F_HRST2	-1,73	.164	.50
Gesamt-Gefallenbewertung	F_HRST1	-3,28	.004	.45
	F_HRST2	-0,43	.738	.12

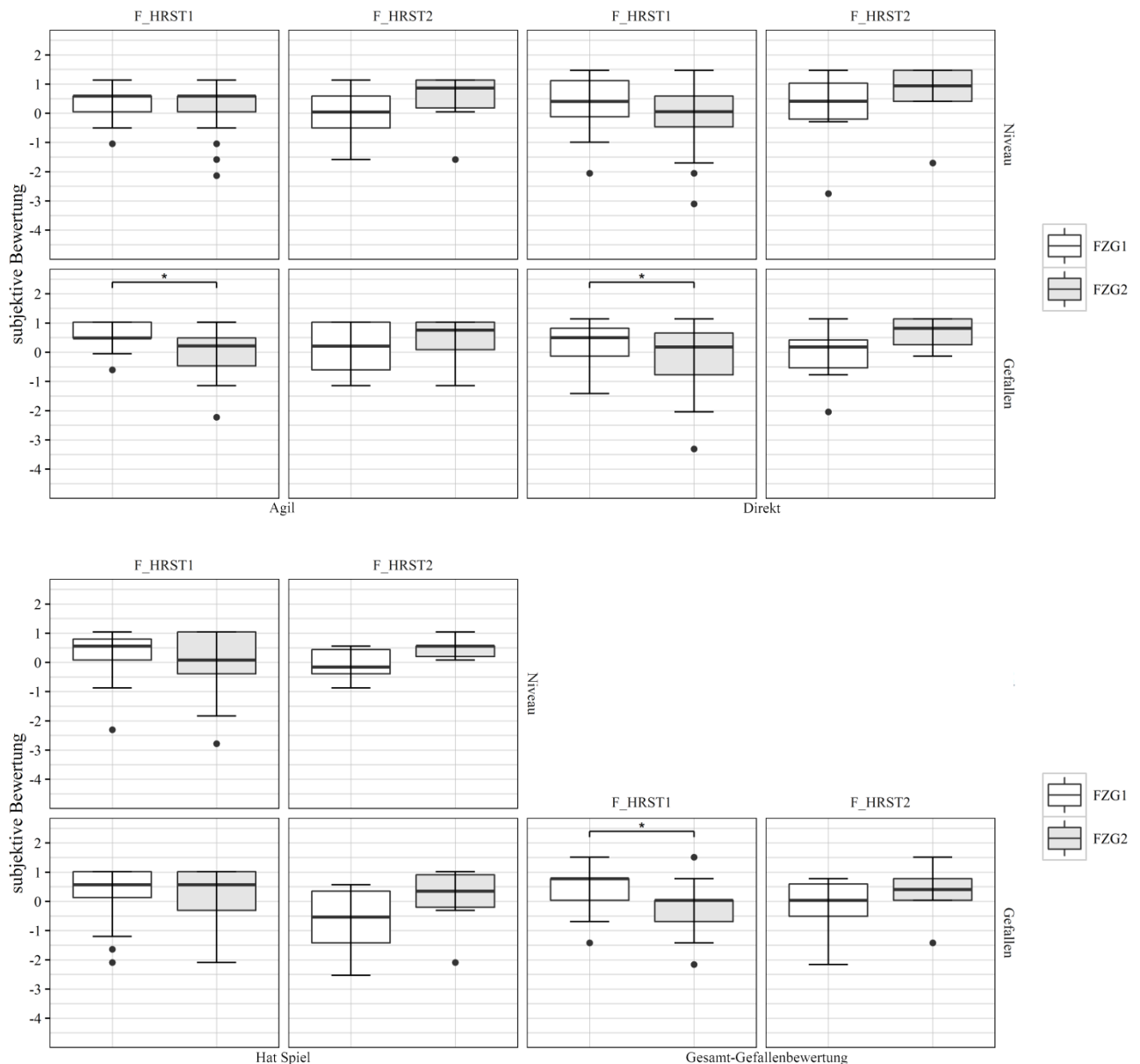


Abbildung 13. Die Mittelwerte der Niveau- und Gefallen-Bewertungen von *agil*, *direkt*, *hat Spiel* und der Gesamt-Gefallenbewertung für die Sub-Stichproben F_HRST1 und F_HRST2 getrennt. Die Fehlerbalken repräsentieren die Standardabweichung.

Die Ergebnisse deuten darauf hin, dass die Aussagekraft durch die geringe Stichprobengröße der Sub-Stichprobe F_HRST2 minimiert wurde. Deshalb wurden die Bewertungen obiger fünf Kriterien sowie fünf weiterer relevanter Bewertungskriterien (*empfindlich*, *stabil*, *leichtgängig*, *schwammig*, *verzögert*) gemeinsam ausgewertet. Hierzu wurden die Niveau- und Gefallen-Bewertungen der beiden Fahrzeuge getrennt analysiert. In Tabelle 21 und Abbildung 14 sind die Ergebnisse der *t*-Tests für unabhängige Stichproben zu finden.

Die Niveau-Bewertungen für FZG1 waren nicht signifikant verschieden zwischen den Sub-Stichproben. Die Gefallen-Bewertungen für FZG1 der Sub-Stichprobe F_HRST1 waren signifikant höher als die der Sub-Stichprobe F_HRST2. Die Niveau- und Gefallen-Bewertungen für FZG2 der Sub-Stichprobe F_HRST2 waren signifikant höher als die der Sub-Stichprobe F_HRST1.

Tabelle 21. Ergebnisse der *t*-Tests für unabhängige Stichproben zum Vergleich aller Niveau- und Gefallen-Bewertungen der beiden Fahrzeuge in Abhängigkeit der Sub-Stichproben F_HRST1 und F_HRST2.

Fahrzeug	Bewertungen	<i>t</i> (327)	<i>p</i>	<i>r</i>
FZG1	Niveau	1,57	.210	.18
	Gefallen	2,59	.032	.29
FZG2	Niveau	-3,58	.002	.33
	Gefallen	-4,87	<.001	.40

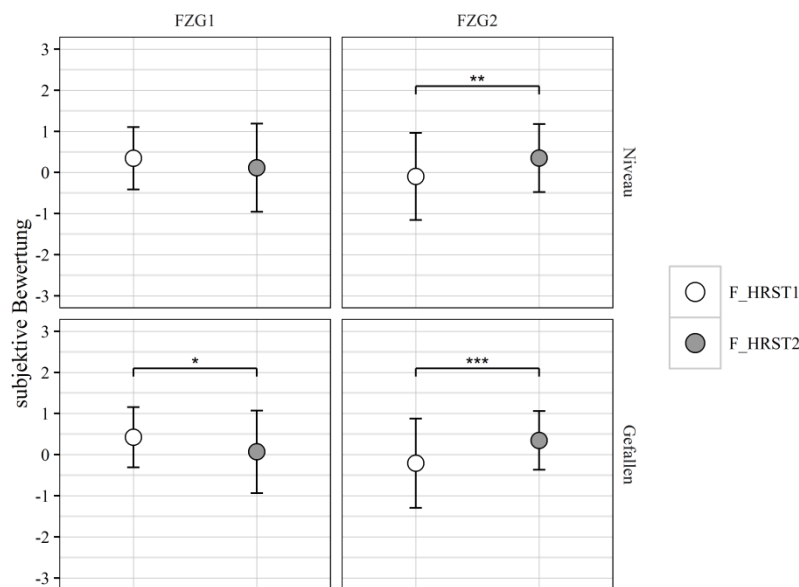


Abbildung 14. Die Mittelwerte aller Niveau- und Gefallen-Bewertungen der beiden Fahrzeuge in Abhängigkeit der Sub-Stichproben F_HRST1 und F_HRST2. Die Fehlerbalken repräsentieren die Standardabweichung.

Zusammenfassend ging aus der Analyse der einzelnen Bewertungskriterien hervor, dass die zuvor angenommenen Unterschiede im Lenkgefühl nicht aufgedeckt werden konnten. Vielmehr traten Unterschiede in entgegengesetzter Richtung auf. Diese lagen jeweils zugunsten des Fahrzeugs vor, dessen Hersteller die Probanden üblicherweise fahren. Die gemeinsame Analyse aller Niveau- bzw. Gefallen-Bewertungen bestätigte diesen Trend. Somit wird die Hypothese, dass die Niveau- und Gefallen-Bewertungen der Fahrzeuge unabhängig davon sind, welches Fahrzeug die Probanden üblicherweise fahren, verworfen.

Hypothese 2: Unabhängigkeit der Bewertungen von der Sub-Stichprobe

H2: Die Bewertungen der Normalfahrer und Experten sind nicht signifikant verschieden.

Für diese Hypothese wurden alle Normalfahrer (F_HRST1, F_HRST2 und F_HRST3) gemeinsam analysiert. Wie bei Hypothese 1 wurden die Niveau- und Gefallen-Bewertungen der zehn Kriterien sowie die Gesamt-Gefallenbewertung herangezogen. Die Auswertung wurde getrennt nach Fahrzeugen sowie für die Niveau- bzw.

Gefallen-Bewertung vorgenommen. Es wurden t -Tests für unabhängige Stichproben berechnet. Die deskriptiven Statistiken sind in Anhang II zu finden.

Für FZG1 waren die Niveau-Bewertungen der Normalfahrer signifikant höher als die der Experten, $t(133,06) = -7,68, p < .001, r = .32$ (Abbildung 15). Die Gefallen-Bewertungen der Normalfahrer waren ebenfalls signifikant höher als die der Experten, $t(134,87) = -9,46, p < .001, r = .37$. Für FZG2 wurden keine Unterschiede der Niveau-Bewertung festgestellt, $t(507) = -1,18, p = .935, r = .05$. Dasselbe galt für die Gefallen-Bewertungen, $t(558) = 0,56, p = .655, r = .02$.

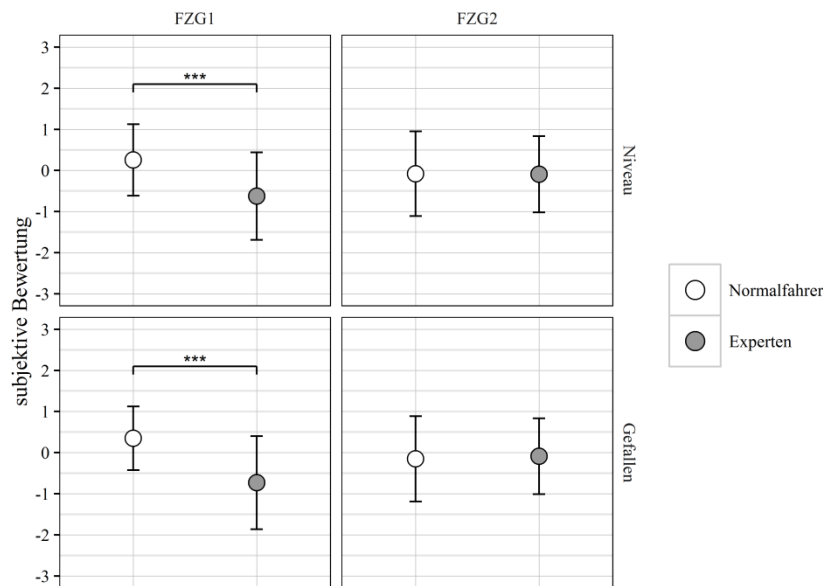


Abbildung 15. Die Mittelwerte aller Niveau- und Gefallen-Bewertungen der beiden Fahrzeuge in Abhängigkeit der Sub-Stichproben Normalfahrer und Experten. Die Fehlerbalken repräsentieren die Standardabweichung.

Somit wurde die Hypothese, dass die Bewertungen des Lenkgefühls der Normalfahrer und die der Experten nicht signifikant verschieden sind, verworfen. Die Daten wurden einer explorativen Untersuchung unterzogen.

Hypothese 3: Konsistenz der Bewertungen der beiden Sub-Stichproben

H3: Die Bewertungen der Experten sind konsistenter als die der Normalfahrer.

Für diese Hypothese wurden alle Normalfahrer (F_HRST1, F_HRST2 und F_HRST3) gemeinsam analysiert. Wie bei Hypothese 1 wurden die Niveau- und Gefallen-Bewertungen der zehn Kriterien sowie die Gesamt-Gefallenbewertung herangezogen. Somit ergaben sich 21 analysierte Variablen. Es wurden t -Tests für unabhängige Stichproben berechnet. Die deskriptiven Statistiken sind in Anhang III zu finden.

H3a: Die wahrgenommenen Unterschiede fallen bei den Experten größer aus als bei den Normalfahrern.

Jeweils für Normalfahrer und Experten getrennt wurden die Bewertungen der FZG1 und FZG2 mithilfe von t -Tests für verbundene Stichproben verglichen. Diese Berechnungen dienten jedoch nur dazu, die Effektstärken r zu bestimmen. Die anschließende Analyse ergab, dass die Effektstärken der Experten signifikant größer waren als die der Normalfahrer, $t(40) = -4,10, p = .001, r = .54$.

Somit wurde die Hypothese, dass die aus dem Vergleich der Beurteilungen der FZG1 und FZG2 berechneten Effektstärken bei den Experten größer ausfallen als bei Normalfahrern, beibehalten.

H3b: Die Beurteilungen der Experten weisen kleinere Standardabweichungen auf als die der Normalfahrer.

Jeweils für Normalfahrer und Experten getrennt wurden die Standardabweichungen SD der Bewertungen zu den 21 Variablen berechnet. Hierzu wurden die entsprechenden Rohwerte beider Fahrzeuge herangezogen. Die Standardabweichungen der Bewertungen der Experten unterschieden sich nicht von denen der Normalfahrer, $t(40) = 0,81, p = .501, r = .13$ (Abbildung 16).

Die Hypothese, dass die Beurteilungen der Experten kleinere Standardabweichungen aufweisen als die der Normalfahrer, wurde verworfen.

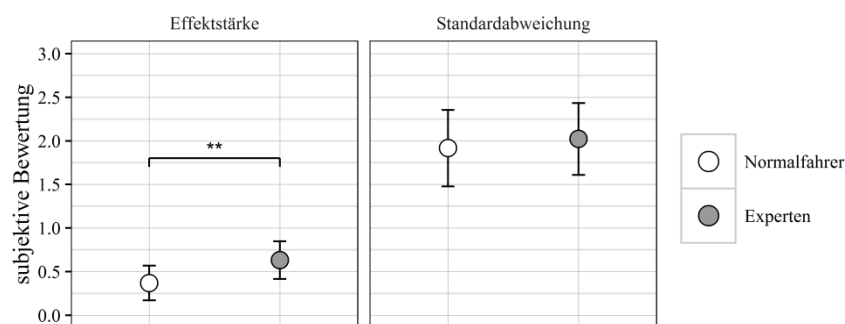


Abbildung 16. Die Mittelwerte der Effektstärke r und Standardabweichung SD in Abhängigkeit der Sub-Stichproben Normalfahrer und Experten. Die Fehlerbalken repräsentieren die Standardabweichung.

Insgesamt wurde die Hypothese, dass die Bewertungen der Experten konsistenter sind als die der Normalfahrer, verworfen.

Explorative Analyse

Explorative Analyse: Wird die Gesamt-Gefallenbewertung der beiden Fahrzeuge durch die Sub-Stichproben Normalfahrer und Experten beeinflusst?

In Hypothese 2 zeigte sich, dass die Beurteilungen der Normalfahrer und Experten signifikant verschieden waren. Während FZG1 bei den Normalfahrern hohe Werte aufwies, war dies bei den Experten für FZG2 der Fall. Deshalb wurden mithilfe einer gemischt-faktoriellen Varianzanalyse (GLM 5) die Gesamt-Gefallenbewertungen beider Fahrzeuge in Abhängigkeit der Sub-Stichproben verglichen und auf einen Interaktionseffekt untersucht. Die deskriptiven Statistiken sind in Anhang IV zu finden.

Der Box-Test war signifikant ($p = .003$), was eine Verletzung der Annahme der Gleichheit der Kovarianzen bedeutet. Nach Tabachnick und Fidell (2014) ist dies jedoch nicht problematisch, wenn, wie in diesem Fall, die Gruppe mit der größeren Fallzahl die größeren Varianzen und Kovarianzen produziert. Der Interaktionseffekt war signifikant, $F(1, 49) = 15,54, p = .001$. Dies bedeutet, dass die Gesamt-Gefallenbewertung von den beiden Prädiktoren Fahrzeug und Sub-Stichprobe abhängt (Abbildung 17).

Deshalb wird die Frage, ob die Gesamt-Gefallenbewertung der beiden Fahrzeuge durch die Sub-Stichprobe beeinflusst wird, positiv beantwortet.

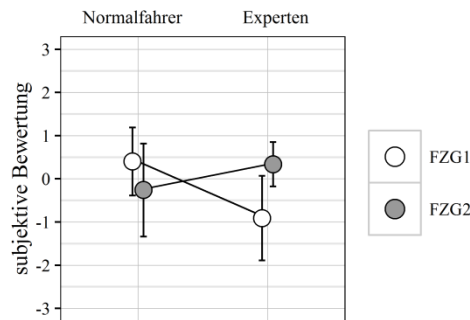


Abbildung 17. Die Mittelwerte der Gesamt-Gefallenbewertung der Sub-Stichproben Normalfahrer und Experten in Abhängigkeit der beiden Fahrzeuge. Die Fehlerbalken repräsentieren die Standardabweichung.

Explorative Faktorenanalyse

Nachfolgend wurde die testtheoretische Güte des Fragebogens beleuchtet. Hierzu wurden die gemittelten, z -transformierten Gefallen-Bewertungen herangezogen. Mithilfe einer explorativen Faktorenanalyse sollten die Zusammenhänge zwischen den Variablen aufgedeckt und diese entsprechend gruppiert werden (Hutcheson & Sofroniou, 2009). Da der Fragebogen für Normalfahrer konzipiert war, wurden die Daten der Experten für diese Analyse nicht verwendet. Aufgrund der zweifachen Beantwortung lagen $N = 82$ Datensätze vor. Dies entspricht der von Kallus (2010) geforderten Stichprobengröße von 50 bis 80 Personen für einen Pretest.

Im Rahmen des vorgeschalteten Datenscreenings wurden die Kaiser-Meyer-Olkin (KMO) Kennwerte berechnet. Ein KMO-Wert nahe 1 bedeutet, dass es Muster von Korrelationen zwischen den Variablen gibt und eine Faktorenanalyse angemessen ist. Der KMO-Wert über alle Items betrug .827 und war gemäß Hutcheson und Sofroniou (2009) als „lobenswert“ einzustufen. Die individuellen KMO-Werte von 28 der 32 Items waren niedriger als der geforderte Wert von .5, die Werte der anderen 4 Items waren kleiner als .6. Ferner wurden die Korrelationen zwischen den Items analysiert. Für sieben Items lag mit weniger als zehn anderen Items eine Korrelation größer als .3 vor. Die ebenfalls berechnete Determinante sollte laut Field (2013) größer als .00001 sein, um Multikollinearität auszuschließen. Im vorliegenden Fall betrug sie 7.164E-11. Bartlett's Test auf Sphärizität wurde angewendet um zu prüfen, ob die Korrelationsmatrix der vorliegenden Daten von der Einheitsmatrix verschieden ist. Der Unterschied war signifikant mit $p < .001$.

Trotz oben ermittelter Schwierigkeiten wurde gemäß Field (2013) mit IBM SPSS 21 eine Hauptkomponentenanalyse durchgeführt. Diese sollte die Komponenten identifizieren, mit denen so viel Varianz wie möglich erklärt werden kann (Kim & Mueller, 1978). Die korrekte Faktoren-Anzahl wurde mit einer Parallel-Analyse bestimmt, welche im Wesentlichen auf dem Vergleich der tatsächlich vorliegenden Daten mit zufällig erzeugten Daten basiert. Sie ist nach Zwick und Velicer (1984) der Analyse des Scree-Plots sowie dem Kaiser-Kriterium vorzuziehen. Bei O'Connor (2000) findet sich eine Syntax zur Durchführung der Parallelanalyse in IBM SPSS. Die Analyse der vorliegenden Daten legte eine einfaktorielle Lösung nahe. Der beizubehaltende Faktor erklärte 35,24 % der Varianz.

Aufgrund oben erläuteter Problematiken wird auf die Darstellung der Faktorstruktur verzichtet. Vielmehr wird abgeleitet, dass dem Fragebogen zu diesem Entwicklungsstand eine unzureichende testtheoretische Qualität zugrunde lag. Um Optimierungsmöglichkeiten aufzudecken, wurden die 40 nicht umgepolten, nicht gemittelten und nicht z -transformierten Niveau-Bewertungen einer weiteren Analyse unterzogen. Hierzu wurden in Anlehnung an Riedel und Arbinger (1997) die prozentualen Häufigkeiten berechnet, wie oft die einzelnen Skalenpunkte ausgewählt wurden (Tabelle 22).

Tabelle 22. Antwortverhalten der Normalfahrer bei nicht-invertierten und invertierten Variablen.

Skalenpunkt	Anteil aller Antworten [%]		Anteil Items an allen Items, bei denen Skalenpunkt von keinem Probanden ausgewählt wurde [%]	
	nicht-invertierte Items	invertierte Items	nicht-invertierte Items	invertierte Items
1	1,00	40,43	65,20	0
2	1,12	20,85	39,10	0
3	2,16	15,43	13,00	5,90
4	3,76	6,93	4,30	5,90
5	4,65	4,16	0	5,90
6	6,44	3,82	0	5,90
7	14,27	3,58	0	23,50
8	21,79	3,52	0	5,90
9	25,57	0,81	0	52,90
10	19,23	0,46	0	70,60

Die Skalenpunkte 1-5 wurden bei den positiv formulierten Items jeweils von weniger als 5 % der Probanden ausgewählt. Die Skalenpunkte am anderen Pol wurden hingegen teilweise von mehr 20 % der Fahrer gewählt. Bei den invertierten Items wurden die Skalenpunkte 5-10 jeweils von weniger als 5 % ausgewählt. Auf den gegenüberliegenden Pol fielen dementsprechend mehr als 40 % der Urteile. Bei den nicht-invertierten Items verteilen sich also die Bewertungen am oberen Drittel der Skala, wobei der Skalenpunkt neben dem Pluspol am häufigsten ausgewählt wurde. Bei den invertierten Items erfolgte eine strikte Ablehnung, welche sich durch die sehr häufige Auswahl des Minuspols selbst zeigte. Es wurde außerdem analysiert, wie oft innerhalb eines Kriteriums einzelne Skalenpunkte nicht ausgewählt wurden. Bei mehr als 2/3 der Kriterien wurde der entgegen dem Konstrukt liegende Pol von keinem Probanden ausgewählt. So wurde den nicht-invertierten Items wie beispielsweise *sicher* und *genau* stark zugestimmt, während invertierte Kriterien wie beispielsweise *nervös* oder *hakt* stark abgelehnt wurden. Auf Basis dieser Erkenntnisse werden in Kapitel 3.2.5 Schritte zur Optimierung abgeleitet und durchgeführt.

3.2.4. Diskussion

Zusammenfassung und Interpretation der Ergebnisse

H1: Die Niveau- und Gefallen-Bewertungen der Fahrzeuge sind unabhängig davon, welches Fahrzeug die Probanden üblicherweise fahren.

Hierzu wurden die Niveau- und Gefallenbewertungen der beiden Fahrzeuge in Abhängigkeit der Sub-Stichproben untersucht. Obwohl auf Basis der herangezogenen Quellen von bestimmten Unterschieden in den Bewertungen auszugehen war, konnten diese nur teilweise aufgedeckt werden. Demgegenüber wurden signifikante Unterschiede gefunden, welche eine Bevorzugung des Fahrzeugs anzeigten, dessen Hersteller die Probanden üblicherweise fahren. Bekräftigt wurde diese Tendenz auch durch die gemeinsame Analyse aller Bewertungen. Die Gefallen-Bewertung von FZG1 sowie die Niveau- und Gefallen-Bewertungen von FZG2 waren signifikant verschieden zwischen den Sub-Stichproben. Somit bevorzugten Probanden, die hauptsächlich einen der beiden Hersteller fuhren, das Versuchsfahrzeug dieses Herstellers. Deshalb wurde Hypothese 1 verworfen.

H2: Die Bewertungen der Normalfahrer und Experten sind nicht signifikant verschieden.

Bei FZG1 waren die Niveau- und Gefallen-Bewertungen der Normalfahrer signifikant höher als die der Experten. Für FZG2 zeigten sich keine signifikanten Unterschiede. Aufgrund der hochsignifikanten Unterschiede und der zugrunde liegenden mittleren Effektstärken wurde Hypothese 2 verworfen. Es ist jedoch hervorzuheben, dass nur bei einem der beiden Fahrzeuge unterschiedliche Bewertungen auftraten.

H3: Die Bewertungen der Experten sind konsistenter als die der Normalfahrer.

H3a: Die wahrgenommenen Unterschiede fallen bei den Experten größer aus als bei den Normalfahrern.

Die zuvor berechneten Effektstärken waren bei den Experten signifikant größer als bei den Normalfahrern. Die Hypothese wurde beibehalten. Das Ergebnis entspricht der in der Literatur zu findenden Annahme, dass Experten Unterschiede eher aufdecken können als Normalfahrer (Nybacka et al., 2014).

H3b: Die Beurteilungen der Experten weisen kleinere Standardabweichungen auf als die der Normalfahrer.

Die Standardabweichungen der Normalfahrer und Experten waren nicht signifikant verschieden. Die Hypothese wurde verworfen. Dieses Ergebnis steht im Widerspruch zu Koch (2010), wonach die Bewertungen von Experten kleinere Streuungen aufweisen.

H3 wurde insgesamt verworfen. Die Ergebnisse belegten nicht, dass die Bewertungen der Experten konsistenter sind als die der Normalfahrer. Dies wird gemeinsam mit den Erkenntnissen zu Hypothese 2 im nächsten Kapitel vor dem Hintergrund der Objektivität der Bewertungen diskutiert.

Explorative Analyse: Wird die Gesamt-Gefallenbewertung der beiden Fahrzeuge durch die Sub-Stichproben Normalfahrer und Experten beeinflusst?

In Ergänzung zu den Ergebnissen zu Hypothese 2 wurde die Gesamt-Gefallenbewertung auf einen Interaktionseffekt mit den beiden Prädiktoren Fahrzeug und Sub-Stichprobe (NF vs. EXP) untersucht. Es wurde ein hoch signifikanter Effekt gefunden, weshalb der Aussage zugestimmt wird.

Insgesamt legte die Diskrepanz zwischen den Bewertungen der Normalfahrer und Experten den Schluss nahe, dass die Ergebnisse nur unter Vorbehalt interpretiert werden sollten. Das Augenmerk soll nun darauf gelegt werden, das Antwortverhalten der Sub-Stichproben zu erklären.

Objektivität und Reliabilität der Bewertungen der beiden Sub-Stichproben

Im folgenden Abschnitt wird diskutiert, inwiefern und wodurch die Objektivität und Reliabilität der Beurteilungen der Normalfahrer und Experten eingeschränkt worden sein könnte.

Im ersten Schritt wird das Antwortverhalten der Normalfahrer beleuchtet. Die Hypothesentests legten nahe, dass die Probanden dasjenige Versuchsfahrzeug besser bewerteten, dessen Hersteller sie auch im täglichen Fahrbetrieb fuhren. Dies widersprach jedoch der Annahme, dass das üblicherweise gefahrene Fahrzeug keinen Einfluss auf die Bewertung hat. Somit konnte mit dem Versuchskonzept keine neutrale Bewertung erreicht werden. Vielmehr bestätigten sich vorherige Ergebnisse, wonach die subjektiven Bewertungen der Probanden durch den Hersteller bzw. die Marke eines Versuchsfahrzeugs beeinflusst werden können (Schalz et al., 2002).

Diese Beeinflussung könnte durch den Ownership-Effekt nach Beggan (1992) verstärkt worden sein. Dieser Effekt besagt, dass Objekte von Kunden generell besser beurteilt werden als von Nicht-Kunden. Dies wird hauptsächlich mit dem größeren Maß an Zuneigung, Vertrautheit, Wissen und Involvement begründet (Buchner, 2008). Weiterhin könnte auch die Gewohnheit vieler Fahrer an die typischen Eigenschaften des Lenkgefühls von Hersteller 1 dazu geführt haben, dass sie die Lenkeigenschaften von Fahrzeug 1 besser beurteilten. Dies wurde eventuell dadurch unterstützt, dass das Lenken automatisiert abläuft und deshalb insgesamt sehr schwer zu beurteilen ist. Da mehr als 2/3 der Normalfahrer regelmäßige Fahrer des Herstellers von Fahrzeug 1 waren, könnten die vorgestellten Effekte zu insgesamt höheren Bewertungen von Fahrzeug 1 geführt haben.

Ferner spielte vielleicht auch der aus der Sozialpsychologie bekannte Halo-Effekt (Thorndike, 1920) eine Rolle. Dieser besagt, dass Menschen von bekannten Eigenschaften eines Objekts auf unbekannte Eigenschaften schließen. Angewendet auf die beiden Versuchsfahrzeuge könnten die Teilnehmer vom ansprechenderen, hochwertigeren Exterieur und Interieur von Fahrzeug 1 auch auf ein positiveres Lenkgefühl geschlossen haben (Gies & Marusic, 2000). Demgegenüber könnte die günstigere Ausstattung von Fahrzeug 2 zu einer Abwertung des Lenkgefühls geführt haben.

Die Bevorzugung von Fahrzeug 1 könnte auch dadurch begründet werden, dass die Studie von seinem Hersteller ausgerichtet wurde. Sie fand auf dessen Werksgelände statt, die Versuchsleiter gehörten dem Hersteller an und auch die Aufwandsentschädigung wurde von diesem ausbezahlt. Somit könnten auch diese Randbedingungen zu einer unbewussten Bevorzugung dieses Versuchsfahrzeugs geführt haben. Dieses Phänomen wird als soziale Erwünschtheit bezeichnet (Turner & Martin, 1984). Es liegen Erkenntnisse vor, dass es bei Befragungen mit direkter Interaktion zwischen Versuchsleiter und Versuchsteilnehmer besonders ausgeprägt ist (Richman, Kiesler, Weisband & Drasgow, 1999). Dies war in der vorliegenden Studie der Fall. Auch eine persönliche oder formelle Verhaltensweise des Versuchsleiters beeinflusst das Antwortverhalten der Teilnehmer (Dijkstra, 1987). Somit kann die soziale Erwünschtheit ein möglicher Einflussfaktor der Bewertungen sein.

Im zweiten Schritt wird das Antwortverhalten der Experten beleuchtet. Einem Großteil der Experten waren die Eigenschaften der beiden Versuchsfahrzeuge vor dem Versuch bereits bekannt, da sie die für die Hypothesen-erstellung genutzten objektiven Vermessungen vornahmen. Ferner waren alle Experten mit dem Lenkgefühl der beiden Hersteller im Allgemeinen vertraut. So konnten sie sich bei ihrer Beurteilung auch auf ihr Hintergrundwissen und vorangegangene Erfahrungen beziehen. Demgegenüber stand den Normalfahrern lediglich ihr unmittelbares Gefühl während des Versuchs zur Verfügung. Die Ergebnisse aus Hypothese 3 zur Konsistenz ihrer Beurteilungen müssen vor diesem Hintergrund gesehen werden.

Somit sind bei beiden Stichproben Einschränkungen bezüglich der Objektivität der Bewertung zu vermuten. Die entsprechenden Anforderungen an die Stichprobe werden in Kapitel 3.2.5 abgeleitet.

Evaluation des Versuchskonzepts

Nachfolgend wird diskutiert, welche Schwierigkeiten in Bezug auf das Versuchskonzept auftraten.

Zuerst soll dabei auf den Fragebogen eingegangen werden. Die Faktorenanalyse ergab, dass der KMO-Wert über alle Items zufriedenstellend, die individuellen KMO-Werte aber nicht akzeptabel waren. Bei Dziuban und Shirkey (1974) finden sich Hinweise, dass der KMO-Wert mit einer größeren Anzahl an Variablen und einer geringeren Anzahl an Faktoren zunimmt. Mit einer einfaktoriellen Lösung bei 32 Variablen könnte dieser Zusammenhang den hohen Gesamt-Wert erklären. Ferner nimmt der KMO-Wert mit insgesamt niedrigen Korrelationen ab (Field, 2013). Die mittlere Korrelation über alle Items betrug in diesem Versuch $r = .309$ und war somit als klein einzustufen. Dies ließ sich eventuell auf das breite Spektrum abgedeckter Aspekte zurückführen. Eine weitere Ursache für niedrige KMO-Werte können auch kleine Stichproben sein (Field, 2013). Zur Bestimmung des benötigten Stichprobenumfangs sind neben der Anzahl an Items pro Faktor auch die Kommunalitäten nach der Extraktion entscheidend (Bühner, 2011; MacCallum, Widaman, Zhang & Hong, 1999). Die Kommunalitäten beschreiben den Anteil der Varianz eines Items, welcher durch die beibehaltenen Faktoren erklärt werden kann (Field, 2013). Mit niedrigeren Kommunalitäten geht eine Vergrößerung der Mindeststichprobengröße einher (Mundfrom, Shaw & Ke, 2005). Im vorliegenden Fall waren die Werte von sieben Items kleiner als .2, sodass sie unter der in der Literatur häufig verwendeten Grenze für niedrige Kommunalitäten lagen. Deswegen konnten die dort berichteten Zusammenhänge nicht auf die vorliegenden Daten übertragen werden (MacCallum et al., 1999; Mundfrom et al., 2005).

Auch aus den Anmerkungen der Probanden während der Fahrt wurden wichtige Hinweise zur Verständlichkeit des Fragebogens gewonnen. Einige Kriterien, besonders solche ohne Erklärung, wurden von Normalfahrern und Experten unterschiedlich interpretiert. Ein Beispiel hierfür ist das Kriterium *empfindlich*. Während es von den Normalfahrern häufig mit *Lenkungsstößigkeit* gleichgesetzt wurde, bezogen es die Experten mitunter auf ein ausgeprägtes Ansprechverhalten. Außerdem bezogen die Normalfahrer das Gesamtfahrzeugverhalten viel stärker in ihre Beurteilung ein. Dies entspricht der Definition des Lenkgefühls im weiteren Sinn (Wolf, 2009).

Ferner beobachteten die Versuchsleiter, dass die Normalfahrer Schwierigkeiten mit den aus Expertenfragebögen übernommenen Kriterien hatten. Obwohl die Probanden immer wieder explizit darum gebeten wurden,

Verständnisschwierigkeiten zu äußern, wurden nur sehr wenige Probleme berichtet. Eventuell wollten die sie Unklarheiten nicht äußern, da sie sich in einer Prüfungssituation wähnten. Vermutlich orientierten sie sich zusätzlich daran, ob ein Kriterium eine positive oder negative Eigenschaft des Lenkgefühls darstellte. Dang et al. (2015) berichten ebenfalls, dass eines der von ihnen verwendeten Kriterien von den Normalfahrern nicht als eindeutig gut oder schlecht identifiziert werden konnte. In der Folge wurden keine Unterschiede zwischen den Varianten wahrgenommen und das Kriterium aus der Analyse ausgeschlossen.

Auch der Versuchsablauf könnte Unsicherheiten gefördert haben. Die Bewertungskriterien waren rasch nacheinander zu beantworten, sodass teilweise lediglich 30 sec zur Bewertung eines Kriteriums zur Verfügung standen. In diesem Zeitraum musste die Vorstellung des Kriteriums sowie die begründete Niveau- und Gefallen-Bewertung erfolgen. Ein ähnlicher Zeitraum zur Bewertung wurde bei Gil Gómez et al. (2015) für Experten als ausreichend befunden, um Aspekte des Fahrverhaltens sicher zu bewerten. Dies ist jedoch für Normalfahrer über die große Anzahl an ihnen unbekannten Kriterien hinweg möglicherweise nicht zutreffend.

Obige Problematiken verdeutlichen die Notwendigkeit einer weiteren Anpassung des Versuchskonzepts.

3.2.5. Erste Optimierung des Versuchskonzepts

Anforderungen an den Versuchsaufbau

Die Ergebnisse der ersten Evaluation weisen auf eine Optimierung der Anforderung zur Stichprobe, zum Bewertungsvorgehen und zum Versuchsdesign hin. Zuerst soll auf die Anforderungen an die Stichprobe eingegangen werden.

Wie oben diskutiert, wurden die Bewertungen der Normalfahrer durch die Hersteller der Versuchsfahrzeuge beziehungsweise den Hersteller ihres normalerweise gefahrenen Lkw beeinflusst. Diese Problematik kann immer auftreten, wenn die Variation des Lenkgefühls durch mehrere Fahrzeuge hervorgerufen wird. Durch eine entsprechende Vorauswahl des Probandenkollektivs kann diese Problematik umgangen werden, bspw. indem nur Fahrer ohne Erfahrung mit allen verwendeten Versuchsfahrzeugen zugelassen werden. Dieses Vorgehen ist jedoch mitunter schwer realisierbar. Im Folgenden wird als Ergänzung zum Entwurf des Versuchskonzept eine Lösungsmöglichkeit entwickelt, wie die Affinität von Fahrern in Bezug auf Hersteller bzw. Fahrzeuge kontrolliert werden kann.

Zum einen könnte diese Affinität der Fahrer mit einem geeigneten Fragebogen erfasst werden. Idealerweise kann dies im Rahmen eines Screenings vor der eigentlichen Rekrutierung des Probanden erfolgen. Falls dies nicht möglich ist, sollten Fahrer, die eine besonders ausgeprägte Affinität zu einem bestimmten Hersteller aufweisen, nachträglich ausgeschlossen werden. Hierbei ist jedoch zu beachten, dass die Bearbeitung dieses Fragebogens vor der Fahrzeugbeurteilung ein gewisses Priming bewirken könnte, wodurch der Proband allein durch die Vorbefragung in seiner Wahrnehmung beeinflusst wird. Deshalb sollte diese Bewertung im Anschluss an die Beurteilungsfahrt erfolgen.

Eine Literaturrecherche zu möglichen Fragebogen zeigte, dass der Fokus der bisherigen Forschung auf der Erfassung der Markenpersönlichkeit lag. Aaker (1997) entwickelte die „Brand Personality Scale“ mit den fünf

Skalen Ernsthaftigkeit, Reiz, Kompetenz, Raffinesse und Robustheit. Die Items dieses umfassenden Fragebogens sind jedoch nicht zur Bewertung von Fahrzeug-Herstellern geeignet. Es bleibt schließlich anzuzweifeln, ob ein schweres Nutzfahrzeug *sentimental* oder *familienorientiert* sein kann. Ein kürzerer Fragebogen zur Einschätzung der Persönlichkeit einer Marke von Sturm (2011) beinhaltet lediglich zehn Items und ist somit deutlich ökonomischer. Auch hier sind die genutzten Eigenschaftsworte nicht zur Beschreibung schwerer Nutzfahrzeuge geeignet.

Nachdem diese Herangehensweise nicht zielführend war, sollte nicht die Einstellung gegenüber dem Hersteller sondern das Fahrzeug insgesamt bewertet werden. So werden alle Erfahrungen, die der Proband während der Erprobung bzw. in der Vergangenheit mit einem Fahrzeug bzw. einem Hersteller gemacht hat, eingeschlossen. Schlussendlich erschien der Akzeptanz-Fragebogen von van der Laan, Heino und de Waard (1997) am geeignetsten für die Bewertung des Gesamtfahrzeugs, obwohl er ursprünglich zur Bewertung von Fahrerassistenzsystemen konzipiert wurde. Ferner ist er mit seinen neun semantischen Differentialen zeitsparend in der Erhebung und Auswertung. Darüber hinaus passen auch die beiden Sub-Skalen *Nützlichkeit* und *Zufriedenheit* zu den Anforderungen an ein schweres Nutzfahrzeug.

Somit ergeben sich bei der Variation des Lenkgefühls unter Verwendung verschiedener Fahrzeuge weitere Anforderungen an das Versuchssetting:

- SP_4: Akquise eines hinsichtlich der verwendeten Fahrzeuge möglichst neutralen Probandenkollektivs
- SP_5: Erfassung der Einstellung der Fahrer zu den verwendeten Fahrzeugen und ggf. Ausschluss aus Analyse

Nachfolgend werden die Anforderungen an das Bewertungsvorgehen diskutiert. Wie angesprochen wurde die Qualität der Bewertungen möglicherweise durch die Kürze des Zeitraums, in dem die einzelnen Kriterien evaluiert werden konnten, vermindert. Deshalb wird das Bewertungsvorgehen um folgende Anforderung ergänzt:

- BV_7: Pro Kriterium mindestens Beantwortungszeitraum von 60 sec verfügbar

Schlussendlich werden auch die Anforderungen an das Versuchsdesign erweitert. Zwar erwies sich das Versuchsdesign mit Messwiederholung als geeignet. Es lässt jedoch keine Rückschlüsse darauf zu, wie groß die wahrgenommenen Unterschiede zwischen Fahrzeugen bzw. Varianten im Vergleich dazu sind, wenn dasselbe Fahrzeug bzw. dieselbe Variante mehrfach bewertet wird. Es kann also keine Aussage getroffen werden, ob gefundene Unterschiede tatsächlich größer sind als das Rauschen, das durch eine mehrfache Bewertung zustande kommt. Außerdem kann mithilfe einer wiederholten Bewertung die Zuverlässigkeit des Vorgehens bestimmt werden, bspw. anhand der Retest-Reliabilität. Deshalb wird folgende Anforderung aufgenommen:

- VD_2: Wiederholte Bewertung von Fahrzeugen bzw. Varianten

Bewertungskriterien

Wie eingangs erwähnt, sollte auch der Fragebogen einer Überarbeitung unterzogen werden. Dies erfolgte im Wesentlichen durch einen Ausschluss von Kriterien sowie der Verbesserung verbliebener Kriterien. Die Bewertungskriterien inkl. Erklärungen und relevanten Fahrsituationen sind in Tabelle 23 zu finden.

In einem ersten Schritt wurden acht Kriterien ausgeschlossen, da sie bereits durch andere, in der Lenkungsauslegung geläufigere Kriterien abgedeckt werden. Ein Beispiel hierfür ist *spurgetreu*, welches inhaltlich bereits durch *Korrekturbedarf* abgedeckt wurde. Im zweiten Schritt sollten die verbleibenden Kriterien weiterentwickelt werden. Im Zuge dessen wurde ein Kriterium aus zwei ausgeschlossenen Kriterien neu entwickelt, nämlich *Fahrbahnkontakt* aus *empfindlich* und *schwimmen*. Dies war nötig, da bisher der positive Aspekt eines ausgeprägten Fahrbahnkontakts, nämlich eine angenehme Rückmeldung von der Straße, nicht abgebildet wurde. Weitere zwei Kriterien wurden in jeweils zwei neue Kriterien aufgeteilt, z. B. Rückstellverhalten in *Rückstellverhalten-Rückstellgeschwindigkeit* und *Rückstellverhalten-Restlenkwinkel*. Entsprechende Erklärungen für diese fünf neuen Kriterien wurden entwickelt.

Tabelle 23. Auswahl der Bewertungskriterien mit dazugehöriger Erklärung und Fahrsituation.

Nr.	Bewertungskriterium*	Erklärung	Spezifische Fahrsituation
1	Agil	Die Fahrzeugreaktion tritt sofort nach Ihren Lenkeingaben ein.	Mittlere Lenkwinkel in beide Richtungen bei mittleren bis hohen Geschwindigkeiten
2	Anlenkmoment	Zum Anlenken ist ein geringer Kraftaufwand nötig. Hinweis: Der Fokus liegt auf dem Auslenken aus der Mitte.	Kleine Lenkwinkel bei höheren Geschwindigkeiten
3	Direkt	Um eine bestimmte Fahrzeugreaktion zu erreichen, ist nur ein geringer Lenkwinkel nötig.	Alle Lenkwinkelbereiche in allen Geschwindigkeitsbereichen
4	Fahrbahnkontakt	Die Lenkung liefert eine angenehme Rückmeldung über Zustand, Griffigkeit und Anregungen der Straße.	Mittlere und hohe Geschwindigkeiten
5	Flattert (i)	Das Lenkrad zappelt (Links-rechts Bewegungen, vergleichbar mit schlecht abgefahrenen Reifen).	Geradeausfahrt bei hohen Geschwindigkeiten
6	Genau	Das Fahrzeug reagiert genauso, wie Sie es aufgrund Ihrer Lenkeingabe erwarten. Achten Sie darauf, wie gut der von Ihnen gewünschte Abstand zur linken und rechten Fahrbahnmarkierung eingehalten wird.	Engstelle
7	Gleichmäßig	Die Lenkkraft ist über alle Lenkwinkelbereiche gleichmäßig. Hinweis: Der Fokus liegt auf dem Einlenkvorgang aus der Mitte.	Große Lenkwinkel bei niedrigen Geschwindigkeiten
8	Hakt (i)	Beim Einlenken stockt an manchen Stellen die Lenkung, sodass Sie plötzlich mehr Kraft aufwenden müssen („Berg“, der überlenkt werden muss).	Große Lenkwinkel bei niedrigen Geschwindigkeiten
9	Haltemoment	Für das Halten des Lenkrads in Kurven ist ein geringer Kraftaufwand nötig. Hinweis: Der Fokus liegt auf dem eingelenkten Zustand, wenn Sie das Lenkrad nur noch festhalten.	Konstanter Lenkwinkel
10	Hat Spiel (i)	Der von Ihnen eingegebene Lenkradwinkel führt zu keiner spürbaren Fahrzeugreaktion. Es gibt einen toten Bereich um die Mitte.	Geradeausfahrt bei hohen Geschwindigkeiten

11	Knarzt (i)	Sie können ein Lenkgeräusch hören und evtl. fühlen.	Große Lenkwinkel bei niedrigen Geschwindigkeiten
12	Komfortabel	Die Lenkung fühlt sich komfortabel an.	Global
13	Korrekturbedarf (i) - Häufigkeit	Sie müssen häufig gegenlenken, um das Fahrzeug auf der Spur zu halten.	Geradeausfahrt / kleine Lenkwinkel bei mittleren Geschwindigkeiten / Querneigung
14	Korrekturbedarf (i) - Lenkwinkel	Sie benötigen große Lenkwinkel, um das Fahrzeug auf der Spur zu halten.	Geradeausfahrt / kleine Lenkwinkel bei mittleren Geschwindigkeiten / Querneigung
15	Leichtgängig	Um zu lenken ist nur ein geringer Kraftaufwand notwendig.	Große Lenkwinkel bei niedrigen Geschwindigkeiten
16	Lenkaktivität (i)	Sie müssen beim Lenken viel kurbeln.	Große Lenkwinkel in beide Richtungen bei niedrigen Geschwindigkeiten
17	Lenkungsstößigkeit (i)	Das Lenkrad leitet Störungen der Straße weiter. D. h. Sie können einen Schlag im Lenkrad spüren.	Schlechte Fahrbahnoberfläche (Kanaldeckel, ausgeprägte Teerfugen)
18	Mittenzentrierung	Die Geradeausstellung des Lenkrads ist deutlich spürbar. D. h. Sie können die Mitte fühlen, wenn Sie leicht hin und her lenken. Das Lenkrad „fällt“ in die Mitte.	Geradeausfahrt bei mittleren und hohen Geschwindigkeiten
19	Nervös (i)	Das Fahrzeug reagiert hektisch und stärker, als Sie es erwartet haben.	Geradeausfahrt bei hohen Geschwindigkeiten
20	Rückstellverhalten - Geschwindigkeit	Das Lenkrad dreht schnell zurück.	Nach großen Lenkwinkeln bei niedrigen Geschwindigkeiten
21	Rückstellverhalten - Restlenkwinkel (i)	Sie müssen aufgrund eines großen Restlenkwinkels viel nachlenken.	Nach großen Lenkwinkeln bei niedrigen Geschwindigkeiten
22	Sicher	Die Lenkung vermittelt ein sicheres Gefühl.	Global
23	Synthetisch (i)	Die Lenkung liefert keine Rückmeldung des Fahrzustandes (z. B. Radstellung). Die Rückmeldung am Lenkrad entspricht nicht der Fahrbahn bzw. dem Fahrzustand.	Fahrbahnoberfläche (Fahrbahnmarkierung, Querfuge)
24	Verzögert (i)	Sie haben den Eindruck, dass das Fahrzeug spät auf Ihre Lenkbewegungen reagiert.	Mittlere Lenkwinkel in beide Richtungen bei mittleren bis hohen Geschwindigkeiten
25	Vibriert (i)	Es sind schnelle Schwingungen am Lenkrad spürbar.	Geradeausfahrt bei hohen Geschwindigkeiten

* (i): invertiert

Bei den übrigen Kriterien wurden die Erklärungen zum größten Teil überarbeitet, indem sie verhaltensnaher formuliert oder konkrete Lenkanweisungen eingeschlossen wurden. Beim Kriterium *genau* lautete beispielsweise die neue Erklärung: „Das Fahrzeug reagiert genauso, wie Sie es aufgrund Ihrer Lenkeingabe erwarten. Achten Sie darauf, wie gut der von Ihnen gewünschte Abstand zur linken und rechten Fahrbahnmarkierung eingehalten wird.“ Besonderes Augenmerk wurde außerdem auf den Korrekturbedarf gelegt, da dieser das Lenkgefühl besonders bei längeren Fahrten auf Landstraße oder Autobahn prägt. In den drei Fahrsituationen Geradeausfahrt, langgezogene Kurve und bei Querneigung ist der nötige Korrekturbedarf besonders präsent. Deshalb sollten diese Situationen getrennt und als eigenständige Kriterien erfasst werden.

Bewertungsskala

Neben der Verständlichkeit der Bewertungskriterien wird das Antwortverhalten von Probanden auch stark von der verwendeten Skala beeinflusst. Anhand empirischer Ergebnisse aus der Literatur wird nun hergeleitet, wie die Qualität der Bewertung verbessert werden kann.

Wie in obiger Studie dargelegt, weisen Probanden bei Befragungen zu positiven psychologischen Konstrukten häufig eine ausgeprägte Zustimmungstendenz auf. So wurde in der Forschung zur Resilienz (Friborg, Martinussen & Rosenvinge, 2006) herausgefunden, dass die Verwendung eines semantischen Differentials im Gegensatz zur Likert-Skala Antwortstile reduzieren und zu einer passenderen Faktorstruktur beitragen kann. Obwohl es kognitiv anspruchsvoller zu beantworten ist, kann es insbesondere bei invertierten Items zu einer Verminderung der Fehlantworten führen. Auch der extreme Antwortstil kann durch den Einsatz dieses Skalenformats reduziert werden (Rocereto, Puzakova, Anderson & Kwak, 2011). Die Überlegenheit des semantischen Differentials könnte darin begründet sein, dass sich die Probanden bei dieser Skala mehr Gedanken zu ihrer Antwort machen müssen. Demgegenüber ist bei einer Likert-Skala schlichtweg eine Zustimmung oder Ablehnung nötig.

Nach dieser klaren Empfehlung für das semantische Differential wurde versucht, den Fragebogen entsprechend anzupassen. Dabei zeigte sich jedoch die Problematik, dass nicht für alle Kriterien geeignete Gegensatzpaare gefunden werden konnten. Einige Kriterienpaare, beispielsweise *agil* / *verzögert* oder *ruhig* / *nervös*, schienen vereinbar. Für andere Kriterien wäre lediglich eine Negierung möglich, z. B. *knarzt* / *knarzt nicht*. Diese ist jedoch nach Kallus (2010) nicht zulässig. Die Problematik der inhaltlich korrekten, sinnvollen Gegensatzpaare wird häufig diskutiert (Bianchi, Savardi & Burro, 2011; Schimmack, 2001). Zusammenfassend war also die Bildung von semantischen Differentialen mit inhaltlich belastbaren Gegenpolen für den komplexen Sachverhalt des Lenkgefühls nicht möglich. Somit wurde das Augenmerk auf die Verbesserung der Likert-Skala gelegt.

Zuerst wird auf die optimale Anzahl an Skalenpunkten eingegangen. Diese wurde initial auf Basis der vorliegenden Literatur zum Lenkgefühl auf 10 festgelegt. In der Literatur finden sich methodenorientierte Herangehensweisen zur Festlegung der optimalen Anzahl an Antwortalternativen. Das Ziel ist dabei immer, die Gütekriterien und somit die Datenqualität zu verbessern. Dabei wird häufig eine systematische Variation der Antwortalternativen vorgenommen (Preston & Colman, 2000; Weng, 2004).

Alwin und Krosnick (1991) griffen auf eine sehr große Anzahl bestehender Daten aus Befragungen der Bevölkerung zurück. Sie fanden, dass die Reliabilität von Rating-Skalen zwischen 3 und 7 Stufen kontinuierlich zunahm. Zwischen 7 und 9 Antwortalternativen war jedoch kein Zuwachs mehr zu finden. Eine andere Studie berichtet ab 5 Skalenpunkten zufriedenstellende Werte für Reliabilität und Validität (Preston & Colman, 2000). Die zuverlässigsten Ergebnisse traten hier bei Skalen mit 7 bis 10 Skalenpunkten auf. Bei 10 Skalenpunkten setzte die Sättigung und anschließende Abnahme der Test-Retest-Reliabilität ein. Insgesamt legte diese Studie eine Stagnation der Verbesserungen ab 7 Skalenpunkten nahe. Zusätzlich zur Bestimmung der Gütekriterien wurde eine Nachbefragung der Probanden durchgeführt. Diese ergab, dass sie die Formate mit 7, 9 und 10 Skalenpunkten bevorzugten.

Auch eine weitere Monte-Carlo-Studie mit ähnlicher Zielsetzung fand steigende Reliabilitäten und Validitäten, wenn die Anzahl an Antwortalternativen steigt (Lozano, García-Cueto & Muñiz, 2008). Hier wurden bereits ab 4 Skalenpunkten zufriedenstellende Werte gefunden. Aufgrund der Sättigung beim Zuwachs schien eine Begrenzung auf 7 Skalenpunkte sinnvoll. Deshalb werden Skalen mit 4 bis 7 Antwortalternativen empfohlen. Dieses Intervall geht auch aus einer weiteren Fragebogen-Studie hervor (Muñiz, García-Cueto & Lozano, 2005).

Auch Weng (2004) bestätigt die Mindestanzahl von 5 Skalenpunkten, da bei einer geringeren Anzahl an Antwortalternativen in Abhängigkeit der Stichprobe unterschiedliche Werte der Reliabilität vorlagen. Dies verdeutlicht die Bedeutung der Zielgruppe für die Festlegung der Skala. Während für Studenten Skalen mit 6 bis 7 Antwortalternativen als geeignet angesehen wurden, blieb ihre Eignung für die allgemeine Bevölkerung offen. Diese Einschätzung wird durch andere Belege unterstützt, wonach 7-stufige Skalen für Personen einsetzbar sind, die hoch gebildet sind oder Erfahrung in der subjektiven Bewertung haben (Weijters, Cabooter & Schillewaert, 2010). Für Studien in der Gesamt-Population werden hingegen 5 Skalenpunkte empfohlen, um die Anforderungen an die Probanden möglichst gering zu halten. Dies steht im Widerspruch zu den oben angesprochenen Erkenntnissen von Preston und Colman (2000), wonach die befragten Personen Skalenformate mit einer höheren Anzahl von Antwortalternativen bevorzugten.

Alles in allem wurden für Skalen mit 5 bis 7 Antwortalternativen die höchsten Reliabilitäten und Validitäten gefunden.

Nun soll der Fokus auf der Möglichkeit der verbalen Benennung der Skalenpunkte liegen. Wie bereits in Kapitel 2.3.3 gezeigt, wird dabei häufig zwischen einer Benennung der Pole oder aller Skalenpunkte unterschieden. Oben vorgestellte Studien geben auch hierzu einige Hinweise.

So führt die Benennung aller Skalenpunkte zu zuverlässigeren Ergebnissen als nur die Benennung der Pole (Alwin & Krosnick, 1991). Insbesondere bei vielen Antwortalternativen ist eine höhere Retest-Reliabilität zu verzeichnen (Weng, 2004). Aufgrund der erhöhten Attraktivität der benannten Antwortalternativen zwischen den Polen werden extreme Antwortstile und Fehlantworten reduziert (Weijters et al., 2010). Die Ergebnisse sprechen zusammenfassend für die verbale Benennung aller Skalenpunkte.

Abschließend soll die festzulegende Skala noch einmal vor dem Hintergrund möglicher Antwortstile (Kallus, 2010) beleuchtet werden. Hier wird zuerst auf die Antwortstile eingegangen, die bei den bisherigen Ergebnissen eine Rolle spielten. Zuerst wird die Zustimmungstendenz (Greenleaf, 1992; Moors, 2008) thematisiert. Weijters et al. (2010) zeigten eine Zunahme dieses Antwortstils bei vorhandener Mittelkategorie, wenn alle Antwortalternativen und nicht nur die Endpunkte benannt waren. Demgegenüber fanden andere Autoren, dass die Akquieszenz unabhängig von der Existenz einer Mittelkategorie ist (O'Muircheartaigh, Krosnick & Helic, 2000). Vielmehr war sie bei älteren, weniger gebildeten und weiblichen Probanden stärker ausgeprägt. Aus diesen Ergebnissen lassen sich keine eindeutigen Empfehlungen ableiten, wie die Zustimmungstendenz minimiert werden kann. Dasselbe gilt für den extremen Antwortstil, bei dem die Endpunkte der Skala überproportional häufig ausgewählt werden (Rossi, Gilula & Allenby, 2001). Nach Kieruj und Moors (2010) kann dieser Antwortstil weder durch die Anzahl der Antwortalternativen noch durch das Vorhandensein einer Mittelkategorie reduziert werden.

Neben den Antwortstilen, die bei den bisher gesammelten Daten auftraten, werden Erkenntnisse zu zwei weiteren Antwortstilen vorgestellt. Die Tendenz zur Mitte beschreibt den Trend, die mittlere Antwortalternative bzw. die mittleren Antwortalternativen auszuwählen. Studien zeigten, dass diese Tendenz erst ab 9 Skalenpunkten auftritt und unabhängig vom Vorhandensein einer Mittelkategorie ist (Kieruj & Moors, 2010). Dies spricht also, wie bereits bei einigen zuvor vorgestellten Ergebnissen, nicht gegen die Nutzung einer ungeraden Anzahl an Skalenpunkten. Abschließend wird die sogenannte status quo heuristic angesprochen, wonach Teilnehmer sich auf eine Antwortalternative festlegen und diese unabhängig vom Item-Inhalt wählen. Dies kann zu einer Reduktion der Antwortgenauigkeit führen. Dieser Antwortstil wird maßgeblich von der bereits diskutierten begrenzten Diskriminationsfähigkeit der Probanden hervorgerufen (Weathers, Sharma & Niedrich, 2005).

Obwohl bezüglich der Antwortstile zahlreiche Ergebnisse vorliegen, lassen sich aus ihnen keine konkreten und eindeutigen Empfehlungen für die Skalengestaltung ableiten. Deshalb wird auf die Erkenntnisse zur Reliabilität und Validität zurückgegriffen. Aus diesen Studien ließen sich Hinweise ableiten, dass mindestens 5 Antwortalternativen angeboten werden sollten. Andererseits war bis 7 Skalenpunkte ein weiterer Anstieg der Validität und Reliabilität erkennbar. Diese Anzahl an Antwortalternativen entspricht auch den Erkenntnissen von Saaty und Ozdemir (2003), wonach die Diskriminationsfähigkeit und Auffassungsgabe des Menschen bei 7 begrenzt ist. Bei der Festlegung der Skalenpunkte ist neben den Gütekriterien auch die zukünftige Nutzung zu bedenken. Die Skala soll dazu geeignet sein, das Lenkgefühl in relativen Vergleichen in Bezug auf eine Referenz zu bewerten. Somit muss es den Probanden möglich sein, anzugeben, dass sie keinen Unterschied wahrnahmen. Dies ist bei der Nutzung einer ungeraden Anzahl an Alternativen am besten durch die Auswahl der Mittelkategorie realisierbar. Diese Antwortoption ist besonders relevant, wenn Varianten wiederholt bewertet werden und tatsächlich kein Unterschied vorliegt. Auch aus Sicht der Gütekriterien ist die Mittelkategorie sinnvoll, da ihr Ausschluss zu einer reduzierten Reliabilität (O'Muircheartaigh et al., 2000) und einem Anstieg der Fehlantworten (Weijters et al., 2010) führt. Deshalb war also die Entscheidung zwischen 5 oder 7 Skalenpunkten zu fällen. Wie oben beschrieben legen einige Autoren nahe, bei Studien in der Gesamtbevölkerung maximal 5 Skalenpunkte zu nutzen (Weijters et al., 2010; Weng, 2004). Diese Aussagen beziehen sich jedoch auf Fragebogenstudien, bei

denen Probanden die Skala ohne Hilfe benutzen müssen. Im vorliegenden Fall steht jedoch immer ein Versuchsleiter für Fragen zur Verfügung. Deshalb wird ab nun eine 7-stufige Skala genutzt.

Absolute Bewertung

Die Lenkung zeigt diese Eigenschaft.						
1	2	3	4	5	6	7
gar nicht	kaum	wenig	teilweise	eher	ziemlich	völlig

Mir gefällt, wie stark diese Eigenschaft ausgeprägt ist.						
1	2	3	4	5	6	7
gar nicht	kaum	wenig	teilweise	eher	ziemlich	völlig

Relative Bewertung

Wie stark zeigt das Fahrzeug / die Variante diese Eigenschaft im Vergleich zum vorher bewerteten Fahrzeug / zur vorher bewerteten Variante / zur Referenz?						
-3	-2	-1	0	1	2	3
sehr viel weniger	viel weniger	etwas weniger	gleich	etwas mehr	viel mehr	sehr viel mehr

Wie sehr gefällt Ihnen, wie stark diese Eigenschaft ausgeprägt ist, im Vergleich zum vorher bewerteten Fahrzeug / zur vorher bewerteten Variante / zur Referenz?						
-3	-2	-1	0	1	2	3
sehr viel weniger	viel weniger	etwas weniger	gleich	etwas mehr	viel mehr	sehr viel mehr

Abbildung 18. Vollständig benannte 7-stufige Skalen zur absoluten und relativen Niveau- und Gefallen-Bewertung.

Ferner lässt sich aus den Ergebnissen zur Skalenbenennung ableiten, dass alle Skalenpunkte verbal beschrieben werden sollten. Die vorliegende Literatur wurde deshalb nach Beschreibungen durchsucht, welche für das Raten der vorliegenden Bewertungskriterien geeignet waren. Es wurden jedoch keine passenden deutschen Benennungen gefunden. Für die absolute Bewertung wurden 13 Begriffe aus der Literatur gesammelt (Bass et al., 1974; Bühner, 2011; Kallus, 2010; Moosbrugger & Kelava, 2012). Sie wurden $N = 21$ Personen im Alter von 21 bis 62 Jahren vorgelegt. Die Teilnehmer wurden aufgefordert, die vorgegebenen Begriffe aufsteigend nach ihrer Intensität zu sortieren. Durch eine qualitative Auswertung wurden die folgenden Skalenbenennungen ermittelt

und entsprechend ihrer Intensität aufsteigend gereiht: *gar nicht / kaum / wenig / teilweise / eher / ziemlich / völlig*. Die daraus abgeleitete Skala ist in Abbildung 18 dargestellt. Für die relative Bewertung ergaben sich die folgenden Skalenbenennungen: *sehr viel weniger / viel weniger / etwas weniger / gleich / etwas mehr / viel mehr / sehr viel mehr*.

3.3. Zweite Evaluation des Versuchskonzepts: Bewertung des Lenkgefühls verschiedener Fahrzeuge (Studie II)

3.3.1. Ziele und Hypothesen

Im folgenden Fahrversuch wurde das optimierte Versuchskonzept erneut angewendet und evaluiert. Auch in diesem Versuch wurde eine möglichst einfache Operationalisierung angestrebt, sodass erneut eine Bewertung des Lenkgefühls verschiedener Sattelzüge im öffentlichen Straßenverkehr erfolgte. Aufgrund der angepassten Anforderungen an die Stichprobe wurde in dieser Studie eine Neutralität der Versuchspersonen hinsichtlich der verwendeten Fahrzeuge vorausgesetzt. Die Bewertungen der Fahrer dienten erneut in erster Linie dazu, den Fragebogen und die Anforderungen an den Versuchsaufbau zu evaluieren. Außerdem sollten detailliertere Erkenntnisse zur Konsistenz der Bewertungen generiert werden.

In folgender Studie wurden vier Sattelzugmaschinen verschiedener Hersteller verglichen. Die Fahrzeuge wurden so ausgewählt, dass sie das gesamte Spektrum der auf dem Markt zu findenden Lenkungsauslegungen abbilden. Es war anzunehmen, dass die Fahrer für einige Aspekte des Lenkgefühls Unterschiede zwischen den Fahrzeugen feststellen. Hieraus wurde Hypothese 1 abgeleitet, welche im nächsten Kapitel basierend auf Experteneinschätzungen weiter verfeinert wird:

H1: Die Bewertungen des Lenkgefühls unterscheiden sich zwischen den Fahrzeugen.

In der vorherigen Studie wurden die Bewertungen der Normalfahrer mit denen von Experten verglichen. Es lagen keine deutlichen Unterschiede im Bewertungsverhalten vor, welche für eine größere Genauigkeit oder Zuverlässigkeit der Experten sprachen. Dennoch sollten die Bewertungen der Normalfahrer in der folgenden Studie systematisch bezüglich ihrer Zuverlässigkeit untersucht werden. Dementsprechend lautete H2:

H2: Die Bewertungen der Fahrer sind zuverlässig.

Hierzu wurden die Beurteilungen auf mögliche Veränderungen bei wiederholter Bewertung untersucht. Hypothese 2a lautet somit:

H2a: Die wiederholte Bewertung der Kriterien innerhalb eines Versuchs führt nicht zu signifikant verschiedenen Ergebnissen.

Zum anderen sollte überprüft werden, inwiefern auch Versuche an unterschiedlichen Tagen zu zuverlässigen Ergebnissen führen. Deshalb wurde das Versuchsdesign um den Faktor der Wiederholungsfahrt ergänzt. Somit wurde eine systematische Wiederholung implementiert. Hypothese 2b heißt wie folgt:

H2b: Die wiederholte Bewertung eines Fahrzeugs zu einem späteren Zeitpunkt führt nicht zu signifikant verschiedenen Ergebnissen.

Abschließend sollte eine Aussage dazu gemacht werden, wie ähnlich sich die wiederholten Bewertungen eines Fahrzeugs und die übrigen Fahrzeugbewertungen sind. Hieraus ergibt sich für Hypothese 2c:

H2c: Die ursprüngliche und die wiederholte Bewertung eines Fahrzeugs sind sich ähnlicher als die ursprüngliche Bewertung des Fahrzeugs und die Bewertungen der anderen Fahrzeuge.

3.3.2. Methodik

Versuchsdesign

Einerseits sollten verschiedene Fahrzeuge hinsichtlich des Lenkgefühls verglichen werden. Hieraus resultierte der Faktor *Fahrzeug*, welche die vier Stufen Fahrzeug 1 – Fahrzeug 4 (FZG1 – FZG4) hatte (Tabelle 24) (within-subject). Zur Überprüfung der Stabilität mussten Bewertungen doppelt erfolgen. Hieraus folgte der zweite Faktor *Wiederholungsfahrt*, welcher die zwei Stufen reguläre Fahrt und Wiederholungsfahrt hatte (within-subject). Unter der regulären Fahrt versteht sich die erste Beurteilung eines Fahrzeugs, unter der Wiederholungsfahrt die zweite Beurteilung desselben Fahrzeugs an einem anderen Versuchstag. Somit lag ein zweifaktorielles Versuchsdesign mit Messwiederholung vor. Die abhängigen Variablen waren die Niveau- und Gefallen-Bewertungen der 25 Kriterien (Kapitel 3.2.5) sowie die abschließende Gesamt-Gefallenbewertung. Die Kriterien zum Korrekturbedarf wurden jeweils in allen drei aufgeführten Fahrsituationen bewertet.

Tabelle 24. Versuchsdesign unter Angabe der Stichprobengrößen.

Faktor 2: Wiederholungsfahrt	Faktor 1: Fahrzeug			
	FZG1	FZG2	FZG3	FZG4
Reguläre Fahrt	$n = 10$	$n = 10$	$n = 10$	$n = 10$
Wiederholungsfahrt	$n = 2$	$n = 3$	$n = 3$	$n = 2$

Stichprobe

Gemäß dem Versuchskonzept sollte die benötigte Stichprobengröße mit G*Power 3.1.9.2 (Faul, 2007; Faul, 2009) festgesetzt werden. Die zur Berechnung benötigten erwarteten Effektstärken wurden aus Studie I abgeleitet und lagen durchschnittlich im mittleren Bereich. Für die Überprüfung der H1 zu den Fahrzeugunterschieden sollten Friedmans ANOVAs berechnet werden. Da G*Power diese nicht unterstützt, wurde die Berechnung für eine ANOVA mit Messwiederholung durchgeführt. Um bei vier Messzeitpunkten mit einer mittleren Effektstärke von $f = .25$ und einer Power von $(1-\beta) = .8$ einen tatsächlich vorhandenen Effekt aufzudecken, wurden $N = 24$ Probanden benötigt. Für eine Power von $(1-\beta) = .9$ waren $N = 30$ Probanden nötig. Wie im nächsten Absatz beschrieben wird, wurde eine anfallende Stichprobe genutzt. Diese war in ihren Kapazitäten begrenzt, sodass lediglich $N = 10$ Probanden an der Studie teilnehmen konnten. Somit wurde die geforderte Stichprobengröße nicht erreicht. Da für Hypothese 2 die Daten mehrerer Bewertungskriterien zusammengelegt wurden, führte die Stichprobengröße hier zu einer ausreichend hohen Power.

Wie in Studie I beschrieben, hatten die Hersteller der Versuchsfahrzeuge und die damit verbundenen Assoziationen des Probanden einen Einfluss auf seine Bewertung. Hieraus wurde die Anforderung (SP_4) abgeleitet, ein hinsichtlich der verwendeten Fahrzeuge möglichst neutrales Probandenkollektiv zu akquirieren.

Deswegen wurden für diese Untersuchung markenneutrale Fahrer gesucht, welche mit keinem der verwendeten oder allen Versuchsfahrzeugen vertraut waren. Da die meisten Speditionen lediglich Fahrzeuge eines Herstellers in ihrem Fuhrpark haben und diese zusätzlich häufig aus älteren Generationen stammen, kam diese Bezugsquelle für Probanden nicht in Frage. Deshalb wurde auf den internen Werkverkehr der Daimler AG zurückgegriffen, welcher einen Fuhrpark mit Fahrzeugen aller Wettbewerber besaß. Diese Stichprobe war für den geplanten Versuch in besonderem Maße geeignet. Die Fahrer wechselten wöchentlich das Fahrzeug, wodurch ihnen die Vor- und Nachteile aller Versuchsfahrzeuge aus dem täglichen Fahrbetrieb bekannt waren. Außerdem war das aufwändige Versuchsdesign mit insgesamt fünf Fahrversuchen pro Person aufgrund der leichten Verfügbarkeit der Werksverkehrsfahrer möglich.

Die $N = 10$ männlichen Fahrer des internen Werkverkehrs waren durchschnittlich 52,4 Jahre alt ($SD = 6,8$), hatten eine mittlere Fahrerfahrung von 29,8 Jahren ($SD = 8,9$) und eine jährliche Fahrleistung von 37.000 km ($SD = 5.000$). Somit entsprach die Stichprobe den Anforderungen des Versuchskonzepts. Die Einstellung der Fahrer zu den verwendeten Fahrzeugen wurde gemäß der erweiterten Anforderungen (SP_5) erhoben. Die angeschlossenen sechs Wilcoxon-Vorzeichen-Rang-Tests zeigten keine signifikanten Unterschiede zwischen den vier Versuchsfahrzeugen. Die Fahrer waren den Fahrzeugen gegenüber also neutral eingestellt, wodurch die Anforderung erfüllt wurde.

Wie bereits in der vorherigen Studie war angedacht, dass auch Experten die Fahrzeuge beurteilen. Aus Gründen der Versuchsökonomie wurde auf diese Form der Experteneinschätzung verzichtet. Vielmehr wurden bereits bestehende Einschätzungen dieser Fahrzeuge zur Spezifizierung der Unterhypothesen zu H1 herangezogen. Da Experten aufgrund ihrer Ausbildung dazu in der Lage sind, allgemein gültige Beurteilungen vorzunehmen, sind diese nicht an spezifische Situationen innerhalb eines Fahrversuchs gebunden. Deshalb sind die Ergebnisse verschiedener Beurteilungsfahrten übertragbar.

Variation des Lenkgefühls

Insgesamt standen Fahrzeuge von sieben verschiedenen Herstellern im Fuhrpark des Werkverkehrs zur Verfügung. Alle entsprachen den typischerweise im Fernverkehr eingesetzten Zugkombinationen. Die Prämisse war, dass die ausgewählten Versuchsfahrzeuge das gesamte Spektrum hinsichtlich des Lenkgefühls schlechter bis guter Lenkungen abdecken. Im Gegensatz zum ersten Fahrversuch lagen jedoch weder verwertbare Berichte aus der Presse noch interne Prüfberichte zum Lenkverhalten der Fahrzeuge vor. Somit konnten allein Experten-Einschätzungen genutzt werden. Diese wurden in Form von zwei Workshops generiert, bei denen Fahrzeuge verschiedener Marken bewertet wurden. Die Beurteilung fand dabei sowohl auf öffentlichen Straßen als auch in synthetischen Fahrmanövern auf der Teststrecke statt. Die Beurteilung dauerte pro Fahrzeug ca. 45 min und schloss verschiedene Aspekte des Lenk- und Fahrverhaltens ein.

Der Pool der Versuchsfahrzeuge setzte sich wie folgt zusammen: Das Lenkgefühl von FZG1 wurde insgesamt am besten eingeschätzt. Mit FZG2 wurde das einzige Fahrzeug mit hybrider Lenkung ausgewählt. So wurde evaluiert, wie die Auslegungsfreiräume dieses neuen Lenksystems derzeit genutzt werden. Außerdem konnten Anhaltspunkte zu möglichen Varianten des Lenkgefühls für den anschließenden Fahrversuch gefunden werden. FZG3 rangierte gemäß der Experteneinschätzung im Mittelfeld und FZG4 wies den Experten zufolge das schlechteste Lenkgefühl auf. Um Aspekte des Lenkgefühls gut vergleichen zu können, sollten Einflussfaktoren möglichst kontrolliert werden. Alle Fahrzeuge wiesen einen Gesamt-Kilometerstand von ca. 150.000 km auf. Radstand und Bereifung waren vergleichbar. Alle Zugfahrzeuge waren mit voll beladenen Euro-Aufliegern gekoppelt.

Im nächsten Schritt wurde Hypothese 1 basierend auf den Experteneinschätzungen weiter detailliert.

H1: Die Bewertungen des Lenkgefühls unterscheiden sich zwischen den Fahrzeugen.

Fahrzeug 2 war als einziges Versuchsfahrzeug bereits mit einem hybriden Lenksystem ausgestattet. Dieses zeichnet sich vor allem bei langsamen Geschwindigkeiten durch eine deutlich größere Lenkunterstützung aus. Deshalb wurde für H1a angenommen:

H1a: Fahrzeug 2 wird hinsichtlich bestimmter Bewertungskriterien besser bewertet als die übrigen Fahrzeuge.

Fahrzeug 1 wurde von den Experten als Benchmark eingeschätzt. Nichtsdestotrotz weist es im Vergleich zu den anderen Fahrzeugen eine ausgeprägte Lenkungsstößigkeit auf. Deshalb gilt für H1b:

H1b: Fahrzeug 1 wird hinsichtlich der Lenkungsstößigkeit schlechter bewertet als die übrigen Fahrzeuge.

Wie zuvor erwähnt, sollten die vier Versuchsfahrzeuge das gesamte Spektrum von Lenkungsauslegungen abbilden. Deshalb sollten ihre Gesamtbewertungen miteinander verglichen werden. So gilt für H1c:

H1c: Die Gesamt-Gefallenbewertungen der Fahrzeuge unterscheiden sich.

Versuchsumgebung und Festlegen der Fahrsituationen

Die Studie wurde auf der Strecke, die die Probanden täglich zurücklegen, durchgeführt. Die Strecke war aus zwei Gründen gut geeignet: Da die Probanden mit ihr vertraut waren, konnten sie sich noch stärker auf die Bewertung des Lenkgefühls konzentrieren. Andererseits bestand die Strecke zum größten Teil aus Autobahnabschnitten und ausgebauten Bundesstraßen, sodass sie den Fahralltag im Fern- und Verteilerverkehr gut abbildete. Die Hin- und Rückfahrt waren insgesamt 120 km lang. Zwischen der Hin- und Rückfahrt lag organisatorisch bedingt eine Pause von ca. 1 Stunde. Diese Pause war der Konzentration dienlich, sodass von einer gleichbleibenden Qualität der Bewertungen auch gegen Ende der Rückfahrt auszugehen war.

Gemäß dem Versuchskonzept wurden die Bewertungskriterien geeigneten Fahrsituationen zugeordnet (Abbildung 19). Jedes Kriterium wurde einmal auf der Hin- und einmal auf der Rückfahrt eingeplant. Neben der Überprüfung von Hypothese 2 hatte dieses Vorgehen den Vorteil, dass durch eine Mittelwertbildung beider Bewertungen genauere Messwerte gewonnen werden. Außerdem konnte beispielsweise bei staubedingten Ausfällen einer Beurteilung zumindest auf die andere Messung zurückgegriffen werden.

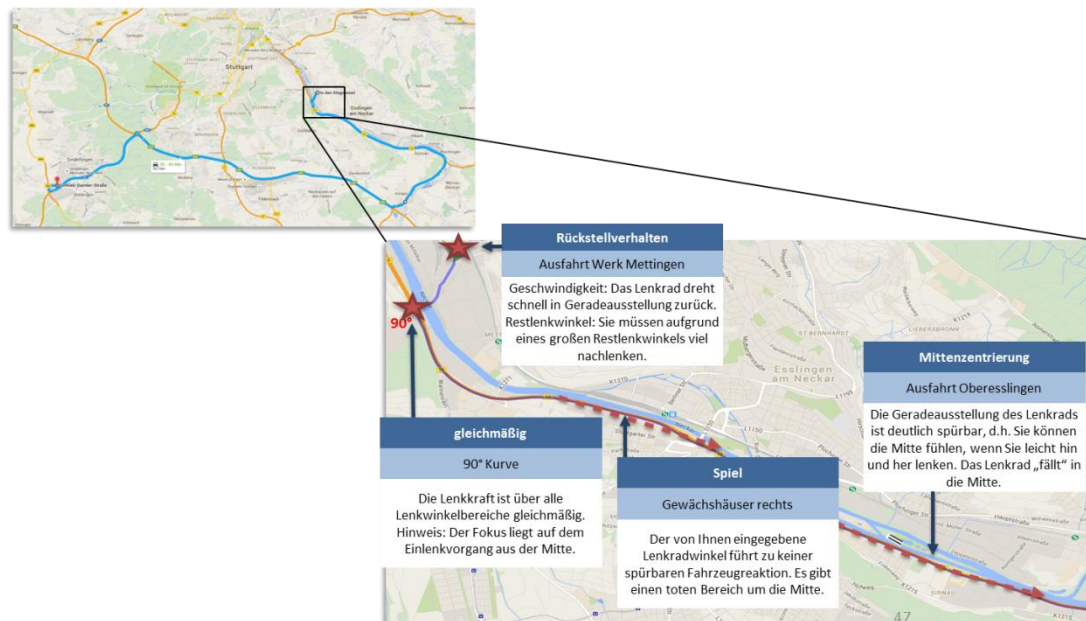


Abbildung 19. Versuchsstrecke mit beispielhaftem Streckenabschnitt zur Verdeutlichung des Bewertungsvorgehens (Kartenmaterial: Google Maps, 2016).

Art der Bewertung, Bewertungsskala und Bewertungskriterien

Gemäß Kapitel 2.3.2 war bei der Kombination aus der gewählten Variation des Lenkgefühls und der Versuchsumgebung eine absolute Bewertung angezeigt. Dementsprechend wurde zur Bewertung eine 7-stufige Absolut-Skala genutzt. Da mit diesem Versuch das gesamte Lenkgefühl evaluiert werden sollte, wurden alle Bewertungskriterien des Fragebogens genutzt. Sowohl der Fragebogen, als auch die Skala finden sich in Kapitel 3.2.5.

Versuchsablauf

Die gesamte Datenerhebung nahm etwa 3 h pro Person in Anspruch. Zuerst fand eine Vorbesprechung statt, um den Probanden einen Einblick in die bevorstehende Beurteilungsaufgabe zu geben. Dazu wurde der Kriterienkatalog inklusive der Erklärungen erläutert. Die Probanden sollten eventuell auftretende Unklarheiten ansprechen, da sich der Fragebogen noch in der Entwicklung befand. Im Anschluss wurde die Bewertung anhand der Aspekte Niveau und Gefallen erläutert und die Beurteilung mit der 7-stufigen Skala an einem Beispiel geübt. Darauf folgte einmalig die Erhebung der demographischen Daten sowie Fahrgewohnheiten. An allen folgenden Versuchstagen fand, falls nötig, eine Auffrischung der Kriterien und der Skala statt.

Anschließend wurde die Hin- und Rückfahrt der oben vorgestellten Strecke absolviert. Die Probanden wurden während der beiden Fahrten von einem Versuchsleiter befragt. Auch jetzt sollten sie auf eventuelle Schwierigkeiten hinweisen. An den zuvor festgelegten Stellen wurden die einzelnen Bewertungskriterien gemäß dem Bewertungsvorgehen beurteilt. Qualitative Anmerkungen des Probanden sowie aufgetretene Verständnisprobleme wurden vom Versuchsleiter notiert. Am Ende jeder Fahrt wurde die Gesamtbewertung in Schulnoten abgegeben und begründet. Abschließend füllten die Fahrer den Akzeptanz-Fragebogen aus. Die Bewertung des nächsten Fahrzeugs fand frühestens am nächsten Tag statt.

Auswertungsvorgehen

Die statistische Auswertung erfolgte mit IBM SPSS Statistics 21. Die Niveau-Bewertungen der invertierten Kriterien wurden umgepolt. Somit standen bei allen Kriterien höhere Niveau-Bewertungen für eine positivere Ausprägung des Lenkgefühls. Die Daten der Niveau- und Gefallen-Bewertungen sowie der Gesamt-Gefallenbewertung wurden dann einer deskriptiven Analyse unterzogen. Da sich die Fahrer erneut in der Nutzung der Skala unterscheiden, wurde eine z -Transformation durchgeführt (Kapitel 3.2.2).

Entsprechend der Voraussetzungen wurden die Daten nur bei Stichprobengrößen von $N < 30$ (Bortz & Schuster, 2010) mithilfe von Kolmogorov-Smirnov-Tests auf Normalverteilung geprüft. Waren die Voraussetzung für parametrische Verfahren nicht gegeben, wurden die entsprechenden nicht-parametrischen Verfahren angewendet.

Das Signifikanzniveau betrug $\alpha = .05$. Signifikanzwerte unter .05 werden mit * gekennzeichnet, Werte unter .01 mit ** und Werte unter .001 mit ***. Entgegen gängiger Auswertungspraxis wurden die aus den Signifikanztests berechneten p -Werte nach Field (2013) nicht halbiert. Zur Vermeidung der α -Fehler-Kumulierung wurde jeweils eine Benjamini-Hochberg-Korrektur für die Hypothesentests bzw. die explorative Analyse durchgeführt (Benjamini & Hochberg, 1995). Für paarweise Vergleiche wurden Effektstärken berechnet. Es gelten gemäß Cohen (1988) für r die Grenzen .1 für einen kleinen, .3 für einen mittleren und .5 für einen großen Effekt.

Die psychometrische Untersuchung des Fragebogens erfolgte mithilfe einer explorativen Faktoren-, Item- und Reliabilitätsanalyse.

3.3.3. Ergebnisse

Hypothese 1: Fahrzeug-Unterschiede

H1: Die Bewertungen des Lenkgefühls unterscheiden sich zwischen den Fahrzeugen.

Um die Unter-Hypothesen H1a - H1c zu überprüfen, wurden die Bewertungen der Fahrer mit Friedmans ANOVAs für verbundene Stichproben untersucht. Bei einem signifikanten Unterschied zwischen den Fahrzeugen wurden Wilcoxon-Vorzeichen-Rang-Tests angeschlossen.

H1a: Fahrzeug 2 wird hinsichtlich bestimmter Bewertungskriterien besser bewertet als die übrigen Fahrzeuge.

Ein Überblick der statistischen Kennwerte ist in Tabelle 25 zu finden. Hinsichtlich der Niveau- und Gefallen-Bewertung von *leichtgängig* beim Rangieren wurden die vier Fahrzeuge nicht signifikant verschieden bewertet (Abbildung 20). Auch für die Bewertungen des *Haltemoments* ergaben sich keine signifikanten Unterschiede.

Tabelle 25. Ergebnisse von Friedmans ANOVAs bzw. Wilcoxon-Vorzeichen-Rang-Tests zum Vergleich der Bewertungen von *Leichtgängig*, *Haltemoment*, *Rückstellverhalten-Geschwindigkeit* und *Rückstellverhalten-Restlenkwinkel*.

Bewertungskriterium	$X^2 (3)$	z	p	r
Leichtgängig (Rangieren) - Niveau	9,80		.050	
Leichtgängig (Rangieren) - Gefallen	7,03		.115	
Haltemoment - Niveau	1,15		.765	
Haltemoment - Gefallen	2,49		.497	
Rückstellverhalten-Geschwindigkeit - Niveau	17,54		.004	
FZG2 vs. FZG1		-1,78	.108	.40
FZG2 vs. FZG3		-1,78	.026	.59
FZG2 vs. FZG4		-1,78	.021	.63
Rückstellverhalten-Geschwindigkeit - Gefallen	4,16		.277	
Rückstellverhalten-Restlenkwinkel - Niveau	11,04		.030	
FZG2 vs. FZG1		-1,368	.212	.31
FZG2 vs. FZG3		-1,368	.029	.56
FZG2 vs. FZG4		-1,368	.031	.57
Rückstellverhalten-Restlenkwinkel – Gefallen	13,67		.018	
FZG2 vs. FZG1		-0,985	.352	.22
FZG2 vs. FZG3		-0,985	.133	.40
FZG2 vs. FZG4		-0,985	.075	.47

Hinsichtlich der Niveau-Bewertung der *Rückstellgeschwindigkeit* waren die vier Fahrzeuge signifikant verschieden. Nachfolgend wurde FZG2 mit den übrigen Fahrzeugen verglichen, wobei sich Unterschiede zu FZG3 und FZG4 zeigten. Die Gefallen-Bewertungen der vier Fahrzeuge hinsichtlich der *Rückstellgeschwindigkeit* waren nicht signifikant verschieden. Die Niveau-Bewertungen des *Restlenkwinkels* beim Rückstellen waren signifikant verschieden. Die Probanden schätzten den verbleibenden Restlenkwinkel von FZG3 und FZG4 größer ein als von FZG2. Die Gefallen-Bewertungen waren ebenfalls signifikant verschieden. Die angeschlossenen Einzelvergleiche ergaben jedoch zu keinem der drei Fahrzeuge signifikante Unterschiede im *Restlenkwinkel*. Die deskriptiven Statistiken sind in Anhang V zu finden.

Die Hypothese, dass Lenkmomente und Rückstellverhalten bei FZG2 mit hybrider Lenkung besser bewertet werden als bei den anderen Fahrzeugen, musste insgesamt verworfen werden. Die vermuteten Unterschiede

konnten nur für die Niveau- und Gefallen-Bewertungen des *Restlenkwinkels* gefunden werden, nicht aber für die anderen Kriterien.

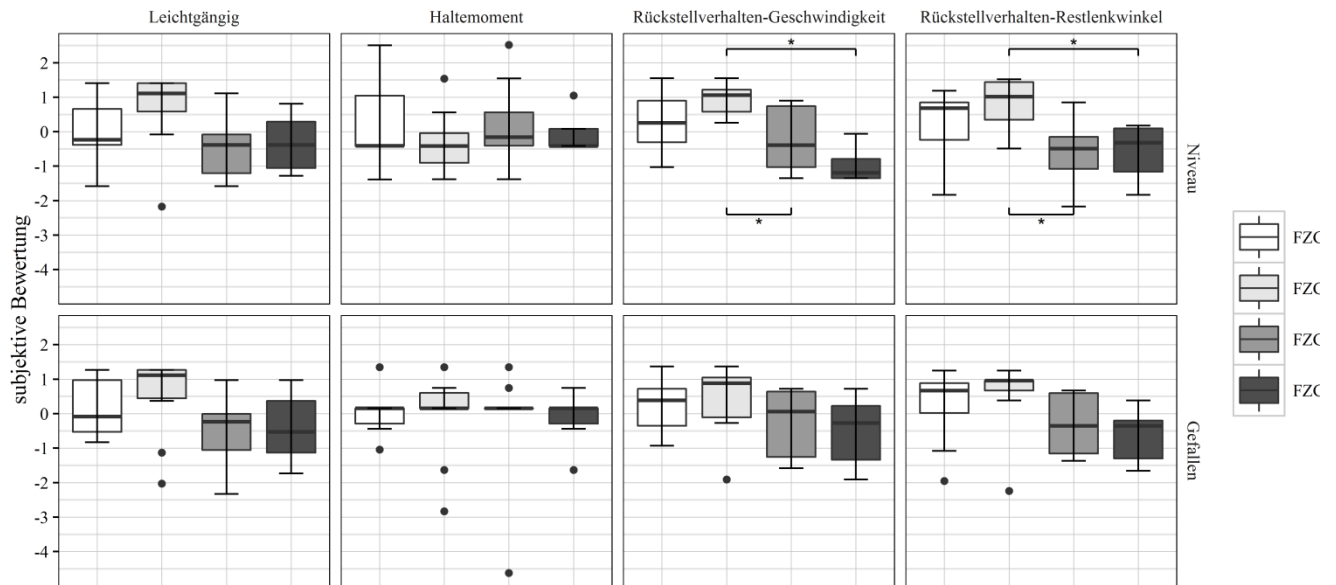


Abbildung 20. Boxplots der Niveau- und Gefallen-Bewertungen von *Leichtgängig*, *Haltemoment*, *Rückstellverhalten-Geschwindigkeit* und *Rückstellverhalten-Restlenkwinkel*.

H1b: Fahrzeug 1 wird hinsichtlich der Lenkungsstößigkeit schlechter bewertet als die übrigen Fahrzeuge.

Bezüglich der *Lenkungsstößigkeit* wurden die vier Fahrzeuge in der Niveau-Bewertung nicht signifikant verschieden eingeschätzt, $\chi^2(3) = 5,29$, $p = .197$ (Abbildung 21). Die Gefallen-Bewertungen waren ebenso nicht signifikant verschieden, $\chi^2(3) = 5,57$, $p = .184$. Die deskriptiven Statistiken sind in Anhang VI zu finden.

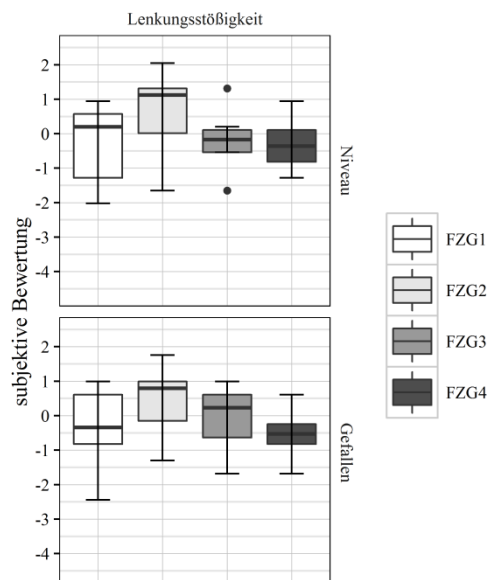


Abbildung 21. Boxplot der Niveau- und Gefallen-Bewertungen von *Lenkungsstößigkeit*.

Somit wurde die Hypothese, dass Fahrzeug 1 im Vergleich zu den anderen Fahrzeugen eine höhere *Lenkungsstößigkeit* und damit niedrigere Gefallen-Bewertungen aufweist, verworfen.

H1c: Die Gesamt-Gefallenbewertungen der Fahrzeuge unterscheiden sich.

Es lag kein signifikanter Unterschied zwischen den Gesamt-Gefallenbewertungen der vier Fahrzeuge vor, $X^2(3) = 7,60$, $p = .096$ (Abbildung 22). Die deskriptiven Statistiken sind in Anhang VII zu finden. Die Hypothese, dass sich die Gesamt-Gefallenbewertungen der vier Fahrzeuge unterscheiden, wurde verworfen.

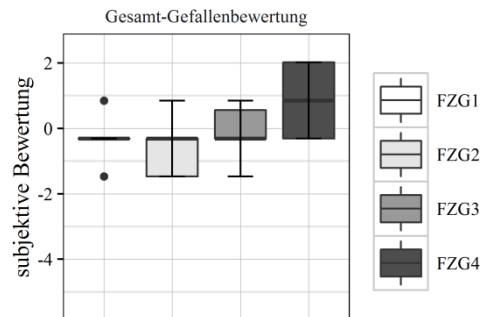


Abbildung 22. Boxplots der Gesamt-Gefallenbewertung.

Aufgrund des Ausbleibens signifikanter Unterschiede wurde Hypothese 1 insgesamt verworfen.

Hypothese 2: Zuverlässigkeit der Beurteilungen

H2: Die Bewertungen der Fahrer sind zuverlässig.

Da eine Analyse auf Kriterien-Ebene eine sehr große Anzahl an Tests gefordert hätte, wurden für diese Auswertung alle Kriterien gemeinsam untersucht. Die Analyse erfolgte dann nach Niveau und Gefallen getrennt. Aufgrund der ausreichend hohen Fallzahlen konnten t -Tests für verbundene Stichproben berechnet werden. Die deskriptiven Statistiken sind in Anhang VIII zu finden.

H2a: Die wiederholte Bewertung der Kriterien innerhalb eines Versuchs führt nicht zu signifikant verschiedenen Ergebnissen.

Der Vergleich der Niveau-Bewertungen ergab keinen signifikanten Unterschied zwischen Hin- und Rückfahrt, $t(1286) = -0,08$, $p = .203$, $r = .00$ (Abbildung 23). Die Gefallen-Bewertungen der Rückfahrt waren ebenfalls nicht signifikant verschieden von denen der Hinfahrt, $t(1402) = 0,15$, $p = .092$, $r = .00$.

Somit wurde die Hypothese, dass die wiederholte Bewertung der Kriterien innerhalb eines Versuchstags nicht zu signifikant verschiedenen Ergebnissen führt, beibehalten.

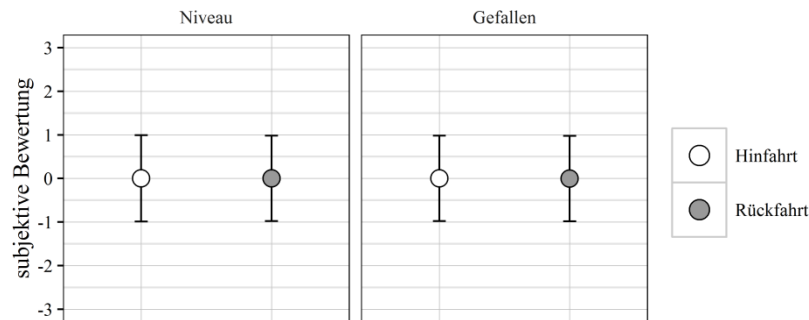


Abbildung 23. Mittelwerte der Niveau- und Gefallen-Bewertungen der Hin- und Rückfahrt. Die Fehlerbalken repräsentieren die Standardabweichung.

H2b: Die wiederholte Bewertung eines Fahrzeugs zu einem späteren Zeitpunkt führt nicht zu signifikant verschiedenen Ergebnissen.

Hierfür wurden die Bewertungen der regulären Fahrt, also der ersten Beurteilung eines Fahrzeugs, mit denen der Wiederholungsfahrt, also der zweiten Beurteilung desselben Fahrzeugs an einem anderen Versuchstag, verglichen (Abbildung 24). Der Vergleich der Niveau-Bewertungen ergab einen signifikanten Anstieg von der regulären Fahrt zur Wiederholungsfahrt, $t(289) = -2,73, p = .025, r = .16$. Derselbe Effekt trat bei den Gefallen-Bewertungen auf, $t(289) = -3,76, p = .001, r = .22$. Die zugrundeliegenden Effektstärken waren klein. Dieses Ergebnis wird in der explorativen Analyse weiter untersucht.

Die Hypothese, dass die wiederholte Bewertung eines Fahrzeugs an unterschiedlichen Versuchstagen nicht zu signifikant verschiedenen Ergebnissen führt, wurde verworfen.

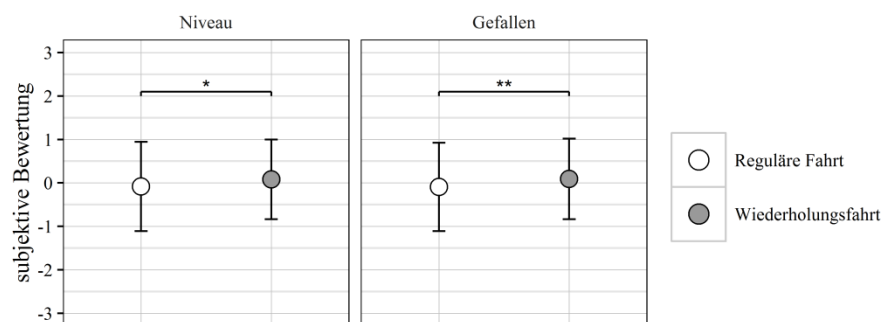


Abbildung 24. Mittelwerte der Niveau- und Gefallen-Bewertungen der regulären Fahrt und der Wiederholungsfahrt. Die Fehlerbalken repräsentieren die Standardabweichung.

H2c: Die ursprüngliche und die wiederholte Bewertung eines Fahrzeugs sind sich ähnlicher als die ursprüngliche Bewertung des Fahrzeugs und die Bewertungen der anderen Fahrzeuge.

Hierfür wurden die Differenzen zwischen den Bewertungen herangezogen. Für alle Kriterien wurden die Differenzen zwischen der regulären Fahrt und der Wiederholungsfahrt des gleichen Fahrzeugs berechnet. Genauso wurden für die vier regulären Fahrten alle sechs Differenzen berechnet und im Anschluss deren Beträge gemittelt. Die gemittelten Differenzen zwischen den Niveau-Bewertungen der vier Fahrten waren signifikant größer als die Differenz zwischen regulärer Fahrt und Wiederholungsfahrt, $t(288) = 13,58, p < .001, r = .62$.

(Abbildung 25). Für die Gefallen-Bewertungen lag ebenfalls ein Unterschied vor, $t(289) = 17,50$, $p < .001$, $r = .72$.

Die Hypothese, dass sich die ursprüngliche und die wiederholte Bewertung eines Fahrzeugs ähnlicher sind als die ursprüngliche Bewertung des Fahrzeugs und die Bewertungen anderer Fahrzeuge, wurde beibehalten.

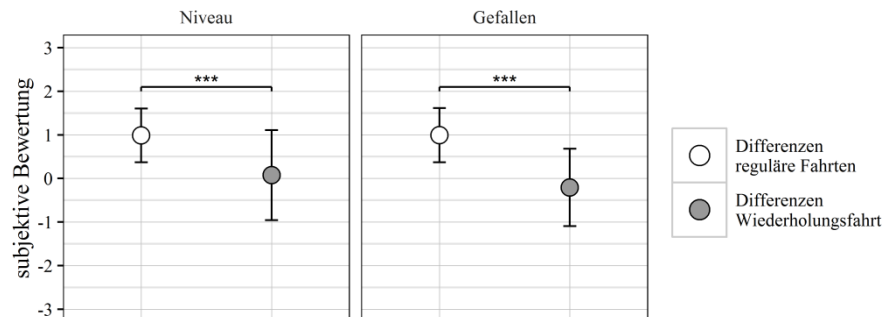


Abbildung 25. Mittelwerte der Differenzen zwischen der regulären Fahrt und der Wiederholungsfahrt sowie zwischen den regulären Fahrten. Die Fehlerbalken repräsentieren die Standardabweichung.

H2 wurde beibehalten, da die Ergebnisse insgesamt für eine Stabilität der Bewertungen der Fahrer sprachen.

Explorative Analyse

EXPL: Die Gefallen-Bewertungen welcher Kriterien unterscheiden sich zwischen regulärer Fahrt und Wiederholungsfahrt?

Die Analysen zu Hypothese H2b zeigten, dass die Gefallen-Bewertungen über alle Kriterien hinweg in der Wiederholungsfahrt signifikant höher waren als in der regulären Fahrt. Mithilfe von Wilcoxon-Vorzeichen-Rang-Tests sollte nun überprüft werden, in welchen Kriterien sich die beiden Fahrten unterscheiden (Tabelle 26). Vor der α -Fehler-Korrektur waren die Gefallen-Bewertungen von drei Kriterien signifikant verschieden, namentlich *Lenkaktivität*, *Mittenzentrierung* und *Korrekturbedarf Häufigkeit (Querneigung)*. In allen drei Fällen wies die Gefallen-Bewertung der Wiederholungsfahrt signifikant höhere Werte auf. Nach der Benjamini-Hochberg-Korrektur war jedoch keiner der Tests mehr signifikant. Die deskriptiven Statistiken sind in Anhang IX zu finden.

Tabelle 26. Ergebnisse der Wilcoxon-Vorzeichen-Rang-Tests für die Kriterien, deren Gefallen-Bewertungen der regulären Fahrt und Wiederholungsfahrt vor der α -Fehler-Korrektur signifikant verschieden waren.

Bewertungskriterium	z	p _{unkorrigiert}	p _{korrigiert}
Lenkaktivität	-2,55	.011	.315
Korrekturbedarf - Häufigkeit (Querneigung)	-2,55	.011	.158
Mittenzentrierung	-2,11	.035	.336

Explorative Faktorenanalyse

Aufgrund der niedrigen Stichprobengröße wurde, wie beim vorherigen Fahrversuch, die Belastbarkeit der Analysen überprüft.

Hierzu wurde eine Hauptkomponentenanalyse mit den gemittelten, z -transformierten Gefallen-Bewertungen berechnet. Die ursprüngliche Extraktion zeigte, dass sechs Faktoren Eigenwerte größer als 1 hatten und insgesamt 78,18 % der Varianz erklärten. Entsprechend der Parallel-Analyse (O'Connor, 2000) sollten zwei Faktoren mit zusammen 59,62 % Varianzaufklärung beibehalten werden.

Auf die in Studie I angesprochene Problematik der ausreichenden Datengrundlage soll zuerst eingegangen werden. Der Großteil der hierzu berechneten Kommunalitäten lag nach der Extraktion zwischen .2 und .8. Ein Wert lag unter .2, 4 Werte über .8. Bei diesen Kommunalitäten und dem vorliegenden Verhältnis von Variablen zu Faktoren war von einer guten bis exzellenten Kongruenz auszugehen (MacCallum et al., 1999; Mundfrom et al., 2005). Unter Kongruenz versteht sich, inwiefern die durch die Stichprobe repräsentierte Gesamtheit die theoretisch definierte Gesamtheit abbildet (Kromrey, 1994). Somit war auf Basis der vorliegenden Datensätze nicht von Einschränkungen der Übertragbarkeit auf die Gesamtpopulation auszugehen. Guadagnoli und Velicer (1988) verfolgen den Ansatz, die notwendige Stichprobengröße mithilfe der Faktorladungen zu bestimmen. Wenn ein Faktor mindestens vier Variablen mit Faktorladungen $> .6$ aufweist, ist er unabhängig von der Stichprobengröße interpretierbar. Dies war in der vorliegenden Analyse der Fall (Tabelle 27). Somit kommen beide Ansätze zu dem Schluss, dass die Interpretation der Faktorenanalyse zulässig war. Weitere Analysen wurden angeschlossen.

Der KMO-Wert von .777 war als „mittelmäßig“ einzustufen (Hutcheson & Sofroniou, 2009). Die individuellen KMO-Werte waren alle kleiner als .5. Ferner wurden die Korrelationen zwischen den Items analysiert. Für zwei Items lag mit weniger als zehn anderen Items eine Korrelation größer als .3 vor. Diese beiden Items, *knarzt* und *flattert*, waren aufgrund ihres Fokus auf Noise, Vibration und Harshness (NVH) auch inhaltlich von den anderen separierbar. Die mittlere Korrelation mit $r = .411$ war höher als im ersten Versuch. Die ebenfalls berechnete Determinante lag mit $1.000E-13$ unter der Grenze von .00001 (Field, 2013), weshalb Multikollinearität erneut nicht ausgeschlossen werden konnte. Der Ausschluss einzelner Items wurde nahegelegt. Bartlett's Test war signifikant mit $p < .001$.

Anschließend wurde eine Rotation durchgeführt (Kim & Mueller, 1978). Da eine Abhängigkeit der Faktoren untereinander anzunehmen war, wurde die oblique Promax-Rotation genutzt (Kline, 2000). Der erste Faktor *Fahrzeugreaktion und Fahrzeugrückmeldung* erklärte mit seinen 18 Items 46,12 % der Varianz. Der zweite Faktor *Lenkkraft* erklärte mit seinen elf Kriterien zusätzlich 13,51 % der Varianz. Die mittlere Korrelation zwischen den Items betrug bei Faktor 1 $r = .560$, bei Faktor 2 $r = .450$.

Tabelle 27. Zweifaktorielle Lösung unter Angabe der internen Konsistenz, Faktorladung und Trennschärfe.

Faktor	Cronbachs α	Bewertungskriterium	Faktorladung	Trennschärfe
Fahrzeugreaktion und Fahrzeug- rückmeldung	.959	Agil	.813	.837
		Direkt	.489	.627
		Fahrbahnkontakt	.370	.462
		Flattert	.284	.220
		Genau	.922	.906
		Hakt	.434	.509
		Komfortabel	.619	.781
		Korrekturbedarf - Häufigkeit (Geradeausfahrt)	.907	.847
		Korrekturbedarf - Lenkwinkel (Geradeausfahrt)	.936	.887
		Korrekturbedarf - Häufigkeit (Kurvenfahrt)	.928	.814
		Korrekturbedarf - Lenkwinkel (Kurvenfahrt)	.974	.890
		Korrekturbedarf - Häufigkeit (Querneigung)	.846	.798
		Korrekturbedarf - Lenkwinkel (Querneigung)	.901	.823
		Mittenzentrierung	.854	.776
		Nervös	.698	.682
		Sicher	.817	.823
		Spiel	.691	.676
		Synthetisch	.645	.642
		Verzögert	.891	.845
Lenkkraft	.891	Anlenkmoment	.820	.702
		Gleichmäßig	.407	.393
		Haltemoment	.832	.612
		Knarzt	.589	.396
		Leichtgängig	1.003	.809
		Lenkaktivität	.676	.713
		Lenkungsstößigkeit	.611	.590
		Rückstellverhalten - Geschwindigkeit	.766	.782
		Rückstellverhalten - Restlenkwinkel	.800	.786
		Vibriert	.486	.544

Weiterhin wurde eine Reliabilitätsanalyse durchgeführt. Unter Reliabilität versteht sich, dass ein Test das gemessene Merkmal exakt und ohne Messfehler misst (Moosbrugger & Kelava, 2012). Zur Bestimmung der internen Konsistenz wurde Cronbachs α berechnet (Cronbach, 1951). Bei Fragebögen mit mehreren Faktoren hängen die Items innerhalb eines Faktors stark zusammen, während die Faktoren untereinander weniger stark korrelieren. Deshalb wurde die interne Konsistenz für beide Faktoren separat berechnet (Bortz & Döring, 2006). Cronbachs α betrug bei Faktor 1 $\alpha = .959$, bei Faktor 2 $\alpha = .891$. In Anlehnung an die bei Cortina (1993) berichteten Werte, die Cronbachs α bei bestimmten Kombinationen von Itemanzahl und mittlerer Korrelation annehmen kann, war bei beiden Faktoren von Eindimensionalität auszugehen.

Auf Basis der gefundenen Faktorstruktur wurde eine Itemanalyse durchgeführt. Hier soll hauptsächlich auf die Trennschärfen eingegangen werden, welche ein Maß für den Zusammenhang zwischen einem Item und dem Gesamtwert sind (Moosbrugger & Kelava, 2012). Bis auf eine Ausnahme waren die Trennschärfen aller Items größer als .3 (Field, 2013). Die Ausnahme war, wie bei den Korrelationen auch, das Item *flattert*.

Anschließend wurde mit den 58 nicht umgepolten, nicht gemittelten und nicht z -transformierten Niveau-Bewertungen eine erneute Analyse der Antwortverteilungen über die 7 Skalenpunkte durchgeführt. Diese wird in Bezug zu den Ergebnissen der ersten Evaluation gesetzt (Kapitel 3.2.3). Da die Studien nicht auf den gleichen Items und den gleichen Fahrzeugen als Bewertungsgrundlage aufbauten, konnte kein unmittelbarer Vergleich durchgeführt werden.

Wie in Tabelle 28 gezeigt, wurden die Skalenpunkte 1-4 bei den positiv formulierten Items jeweils von weniger als 10 % der Probanden ausgewählt. Die Skalenpunkte am anderen Pol wurden hingegen teilweise von mehr als 40 % der Fahrer gewählt. Bei den invertierten Items wurden die Skalenpunkte 4-7 jeweils von weniger als 10 % ausgewählt. Auf die Skalenpunkte am gegenüberliegenden Pol fielen dementsprechend mehr als 40 % der Urteile. Wie im vorherigen Versuch verteilten sich also bei den nicht-invertierten Items die Bewertungen auf der oberen Hälfte der Skala, wobei der Skalenpunkt neben dem Pluspol am häufigsten ausgewählt wurde. Bei den invertierten Items erfolgte ebenfalls eine Ablehnung, wobei sich diese hier durch die Auswahl des Skalenpunkts neben dem Minuspol zeigte. Anschließend wurde analysiert, bei wie vielen Items welche Skalenpunkte nie ausgewählt wurden. Bei mehr als 58 % der nicht-invertierten bzw. 76 % der invertierten Items wurde jeweils der entgegen dem Konstrukt liegende Pol von keinem Probanden ausgewählt. Im vorherigen Versuch betrugen diese Werte 65 % bzw. 70 %. Insgesamt zeigte die Verteilung über die Skalenpunkte immer noch eine starke Zustimmungstendenz auf, die Anzahl „leerer“ Skalenpunkte wurde jedoch reduziert.

Insgesamt waren die Ergebnisse der Hauptkomponenten-, Item- und Reliabilitätsanalyse besser einzustufen als beim vorherigen Versuch. Dies lässt sich zum einen daraus ableiten, dass höhere Interkorrelationen zwischen den Items gefunden wurden. Weiterhin lag eine zweifaktorielle Struktur mit zufriedenstellenden internen Konsistenzen und Trennschärfen vor. Auch der Anteil der erklärten Varianz war höher. Ferner waren die Faktoren inhaltlich erklärbar und wiesen zufriedenstellende Reliabilitäten auf.

Tabelle 28. Antwortverhalten der Probanden bei nicht-invertierten und invertierten Variablen.

Skalenpunkt	Anteil aller Antworten [%]		Anteil Items an allen Items, bei denen Skalenpunkt von keinem Probanden ausgewählt wurde [%]	
	nicht-invertierte Items	invertierte Items	nicht-invertierte Items	invertierte Items
1	1,10	24,92	58,33	0
2	4,59	42,04	12,50	0
3	6,23	11,77	12,50	5,88
4	7,00	8,36	16,67	20,59
5	11,16	6,69	0	17,65
6	49,89	5,57	0	23,53
7	20,02	0,64	0	76,47

3.3.4. Diskussion

Zusammenfassung und Interpretation der Ergebnisse

H1: Die Bewertungen des Lenkgefühls unterscheiden sich zwischen den Fahrzeugen.

H1a: Fahrzeug 2 wird hinsichtlich bestimmter Bewertungskriterien besser bewertet als die übrigen Fahrzeuge.

Fahrzeug 2 mit hybrider Lenkung schnitt nur in einigen Kriterien gegenüber den Fahrzeugen mit hydraulischer Lenkung besser ab. Bei den signifikanten Unterschieden lagen allesamt große Effekte vor. H1a wurde dennoch verworfen.

H1b: Fahrzeug 1 wird hinsichtlich der Lenkungsstößigkeit schlechter bewertet als die übrigen Fahrzeuge.

Weder die Niveau-, noch die Gefallen-Bewertungen waren signifikant verschieden zwischen den vier Fahrzeugen. H1b wurde deshalb verworfen.

H1c: Die Gesamt-Gefallenbewertungen der Fahrzeuge unterscheiden sich.

Die Analysen ergaben keine derartigen Unterschiede in den Gesamt-Gefallenbewertungen der Fahrzeuge, sodass H1c verworfen wurde.

H1 wurde insgesamt verworfen. Es bleibt offen, weshalb die vermuteten Effekte nicht aufgedeckt werden konnten. Dies könnte einerseits in der Operationalisierung begründet sein, da die vorliegende Versuchsumgebung möglicherweise nicht geeignet zur vergleichenden Bewertung des Lenkgefühls war. Dieser Aspekt wird im folgenden Kapitel weiter diskutiert. Andererseits könnte auch das Potential der hybriden Lenkung noch nicht ausgeschöpft sein. Hierfür spricht beispielsweise die extreme Leichtgängigkeit beim Rangieren, welche von zwei Fahrern als zu stark und damit unangenehm empfunden wurde. Das Ausbleiben signifikanter Unterschiede könnte

auch in der niedrigen Stichprobengröße begründet sein. Die Vergleiche waren allesamt nur auf $N = 10$ Bewertungen pro Fahrzeug gestützt, weshalb nichtparametrische Verfahren genutzt werden mussten.

H2: Die Bewertungen der Fahrer sind zuverlässig.

H2a: Die wiederholte Bewertung der Kriterien innerhalb eines Versuchs führt nicht zu signifikant verschiedenen Ergebnissen.

Die Niveau- und Gefallen-Bewertungen der Hinfahrt waren nicht verschieden von denen der Rückfahrt. H2a wurde beibehalten.

H2b: Die wiederholte Bewertung eines Fahrzeugs zu einem späteren Zeitpunkt führt nicht zu signifikant verschiedenen Ergebnissen.

Die wiederholte Bewertung eines Fahrzeugs führte zu signifikant höheren Niveau- und Gefallen-Bewertungen als die erste, reguläre Fahrt. Die Gefallen-Bewertungen wurden in der explorativen Analyse genauer beleuchtet. H2b wurde verworfen.

H2c: Die ursprüngliche und die wiederholte Bewertung eines Fahrzeugs sind sich ähnlicher als die ursprüngliche Bewertung des Fahrzeugs und die Bewertungen anderer Fahrzeuge.

Die untersuchten Differenzen zwischen den Bewertungen der regulären Fahrt und der Wiederholungsfahrt waren kleiner als die zwischen den regulären Fahrten der vier Fahrzeuge. Dies galt sowohl für die Niveau- als auch die Gefallen-Bewertungen. Deshalb wurde H2c beibehalten.

H2 wurde beibehalten, da die Ergebnisse insgesamt für eine Stabilität der Bewertungen der Fahrer sprachen.

EXPL: Die Gefallen-Bewertungen welcher Kriterien unterscheiden sich zwischen regulärer Fahrt und Wiederholungsfahrt?

Bei drei von 29 Kriterien lagen vor der α -Fehler-Korrektur signifikante Erhöhungen der Bewertungen von der regulären Fahrt zur Wiederholungsfahrt vor. Die Kriterien, für die Unterschiede gefunden wurden, beziehen sich auf Aspekte des Lenkwinkels.

Objektivität und Reliabilität der Stichprobe

Im folgenden Abschnitt wird diskutiert, inwiefern und wodurch die Objektivität und Reliabilität der Beurteilungen eingeschränkt worden sein könnte.

Im Gegensatz zur vorherigen Studie ist von einer Objektivität der Bewertungen durch die Stichprobe auszugehen. Diese Markenneutralität der Fahrer wurde anhand eines Fragebogens bestätigt. Zum anderen sprechen hierfür auch die verbalen Anmerkungen der Fahrer während der Befragungen. Diese nannten, abgesehen vom Lenkgefühl, Vor- und Nachteile aller Fahrzeuge. Somit war von keiner systematischen Favoritenbildung über alle Fahrer hinweg auszugehen. Ferner konnten sich die Fahrer durch die Vertrautheit mit den Fahrzeugen und der Versuchsstrecke optimal auf die Beurteilung konzentrieren. Dies wurde dadurch unterstützt, dass die Beurteilungskriterien allen Fahrern spätestens ab der zweiten Fahrt geläufig waren.

Entgegen der Anforderungen des Versuchskonzepts konnte in der vorliegenden Studie die rechnerisch nötige Stichprobengröße nicht erzielt werden. Die geringe Stichprobengröße von $N = 10$ könnte eine Ursache dafür gewesen sein, dass die vermuteten Unterschiede nicht aufgedeckt werden konnten. Somit können diese Ergebnisse

nur mit Vorbehalt interpretiert werden. Andererseits wurden für einige Analysen alle Items gemeinsam ausgewertet, sodass hier deutlich größere Fallzahlen vorlagen. Dies führte andererseits zu einer Überschätzung des Unterschieds zwischen den Gefallen-Bewertungen der Hin- und Rückfahrt, was sich in der vernachlässigbar geringen Effektstärke widerspiegelte.

Diese Ergebnisse bestätigen die ursprünglichen (SP_1 bis SP_3) und die aus der ersten Evaluation gefolgerten (SP_4, SP_5) Anforderungen an die Stichprobe.

Evaluation des Versuchsaufbaus

Bei einer Erprobung im Straßenverkehr ist stets eine Anpassung an die vorgegebenen Verhältnisse nötig. Durch eine große zeitliche Distanz zwischen den einzelnen Bewertungen der verschiedenen Versuchsfahrzeuge in einer bestimmten Situation wird die Vergleichbarkeit zwischen den Fahrzeugen erschwert.

Dies spiegelte sich darin wieder, dass die Gesamturteile über die vier Fahrzeuge hinweg sehr homogen waren. Dies entsprach jedoch in einigen Fällen nicht den Anmerkungen während der Fahrt. So vergaben einige Probanden, obwohl sie ein bestimmtes Fahrzeug während der Fahrt als insgesamt schlechter als ein anderes einschätzten, in der Gesamt-Gefallenbewertung identische Werte. Weiterhin wurde bspw. selten der maximale Skalenwert vergeben. Dies kann dadurch erklärt werden, dass die Spanne der Auslegungen über die Fahrzeuge hinweg nicht überschaubar oder präsent war. Solche Widersprüchlichkeiten traten aber auch bei den übrigen Bewertungskriterien auf. Beispielsweise lagen Diskrepanzen zwischen den vergebenen Bewertungen und den verbalen Begründungen der Probanden vor.

Diese Inkonsistenzen im Antwortverhalten wurden möglicherweise dadurch hervorgerufen, dass die einzelnen Bewertungen zeitlich so weit voneinander getrennt waren. Im Hinblick auf das Aufdecken von Unterschieden im subjektiven Lenkgefühl ist diese Diskrepanz zwischen der quantitativen Bewertung des Probanden und der qualitativen Begründung äußerst problematisch.

Evaluation des Fragebogens

Insgesamt war der in Studie II verwendete Fragebogen auf Basis dieser Ergebnisse als akzeptabel einzuschätzen. Die berichteten Kennwerte müssen vor dem Hintergrund gesehen werden, dass keine psychologischen Konstrukte wie beispielsweise Motivation oder Intelligenz gemessen werden. Somit sollten entsprechende Maßstäbe hinsichtlich der Gütekriterien gesetzt werden. Im Vergleich zur ersten Evaluation waren Verbesserungen hinsichtlich der Qualität sichtbar. Diese spiegelt sich beispielsweise in der höheren erklärten Varianz wider. Die Items des Fragebogens waren zwei inhaltlich erklärbaren Skalen zugeordnet. Die Kriterien *Korrekturbedarf*, *Verzögerung* und *Genauigkeit* korrelierten stark mit dem Faktor *Fahrzeugreaktion und Fahrzeugrückmeldung*. Innerhalb des Faktors *Lenkkraft* wiesen die Aspekte *Leichtgängigkeit*, *Rückstellverhalten* und *Haltemoment* hohe Faktorladungen auf. Daraus lässt sich ableiten, dass sie für das Lenkgefühl von großer Bedeutung sind und auch in zukünftigen Fahrversuchen beachtet werden sollten. Diese 2-faktorielle Lösung stand im Widerspruch zu den in der Literatur zu findenden 3-faktoriellen (Barthenheier, 2004), 4-faktoriellen (Harrer & Pfeffer, 2013) und 8-faktoriellen (Rothhämel et al., 2010) Klassifizierungen des Lenkgefühls. Es ist jedoch anzumerken, dass nur

letztere Klassifizierung auf einem dimensionsreduzierenden Verfahren basiert. Trotz der erzielten Verbesserungen lag eine weitere Reduktion der Bewertungskriterien nahe. Ferner sollten die Bewertungskriterien für eine Bewertung mit relativen Vergleichen optimiert werden.

Ferner wird das Antwortverhalten anhand der 7-stufigen, benannten Skala beleuchtet. Es zeigte sich eine Verringerung der Anzahl an Skalenpunkte, welche von keinem Probanden ausgewählt wurden. Dass trotzdem einzelne Skalenpunkte sehr selten vergeben wurden, kann möglicherweise mit den erprobten Versuchsfahrzeugen begründet werden. Diese deckten bei einigen Kriterien schlichtweg nicht das gesamte Skalenspektrum ab. Am Beispiel der Gefallen-Bewertungen des Kriteriums *Haltemoment* (Abbildung 20) wird deutlich, dass alle Fahrzeuge dieses Kriterium in einem ähnlichen Maß erfüllten. Die geringe Varianz im Antwortverhalten entspricht somit teilweise den tatsächlich vorhandenen Verhältnissen. Wie bereits angesprochen, bestätigt die Häufung bei Skalenpunkt 6 bei nicht-invertierten Items sowie bei Skalenpunkt 2 bei invertierten Items die Äußerungen der Fahrer, wonach sie die Maximalpunktzahl nur ungern vergeben. Es bleibt offen, ob diese Argumentation lediglich für diese spezielle Stichprobe zutrifft oder für alle Normalfahrer.

3.3.5. Zweite Optimierung des Versuchskonzepts

Versuchsdesign

Es zeigte sich, dass die Wiederholung der Bewertungen eine sinnvolle Ergänzung für das Versuchskonzept darstellte. So war es möglich, die Unterschiede zwischen den verschiedenen Fahrzeugen mit dem Rauschen der wiederholten Bewertung eines Fahrzeugs zu vergleichen. Die Wiederholung wurde zum einen dadurch realisiert, dass alle Kriterien sowohl auf der Hin- als auch auf der Rückfahrt bewertet wurden. Ferner wurde ein Fahrzeug zweifach bewertet. Die Anforderung (VD_2), die Bewertung von Fahrzeugen bzw. Varianten zu wiederholen, wurde somit für das finale Versuchskonzept beibehalten.

Zusammenspiel aus Variation des Lenkgefühls, Versuchsumgebung und Art der Bewertung

Die Studien I und II verdeutlichten, dass der Vergleich des Lenkgefühls mit dem vorliegenden Versuchskonzept Probleme mit sich bringt. Wie in Kapitel 2.3.2 (Tabelle 4) dargestellt, kann die Bewertung verschiedener Fahrzeuge bspw. durch unterschiedliche Einstellungen der Probanden oder allein durch den für die Bewertung notwendigen Wechsel zwischen den Fahrzeugen beeinflusst werden. Die Evaluation im Straßenverkehr birgt zusätzlich die Schwierigkeit, dass Fahrsituationen nicht unmittelbar hintereinander mit den verschiedenen Fahrzeugen bzw. Varianten durchfahren werden können. Schlussendlich resultiert aus der Nutzung von verschiedenen Fahrzeugen im öffentlichen Straßenverkehr, dass lediglich absolute Vergleiche realisierbar sind. Diese können jedoch zu den zuvor geschilderten Problemen im Antwortverhalten führen. Die Operationalisierung (OP) ist deshalb insbesondere im Hinblick auf einen möglichst unmittelbaren Vergleich der Varianten bzw. Fahrzeuge auszulegen. Hierzu sind in erster Linie die absoluten Bewertungen durch relative Bewertungen zu ersetzen. Ferner wird die Nutzung einer Referenz angeraten (Kudritzki, 2000; Rothhämel et al., 2011b).

- OP_1: Bewertung erfolgt als relativer Vergleich unter Nutzung einer Referenz

Dies ist jedoch nur durch eine entsprechende Variation des Lenkgefühls und der Versuchsumgebung möglich. Der Zusammenhang zwischen diesen drei Aspekten wurde in Kapitel 2.3.2 bereits dargestellt. Die dort zu findende Tabelle 4 wird nachfolgend noch einmal aufgegriffen (Tabelle 29). Demnach sind sowohl auf dem Testgelände als auch im Fahrsimulator relative Bewertungen realisierbar. Bei Studien im Fahrsimulator ist die Übertragbarkeit der gewonnenen Ergebnisse jedoch maßgeblich von dessen Qualität abhängig. Ferner sind Fahrsimulatoren mit einer entsprechend guten Darstellung häufig schwer verfügbar. Außerdem fallen im Vergleich zur Erprobung im Fahrzeug zusätzliche Aufwände zur Realisierung der Versuche an. Deshalb findet im Rahmen dieses Versuchskonzepts eine Fokussierung auf das Testgelände als Versuchsumgebung statt.

- OP _2: Teststrecke als Versuchsumgebung

Allein die Verlagerung des Versuchs vom Straßenverkehr auf die Teststrecke würde jedoch noch zu keinen wesentlichen Verbesserungen führen. Vielmehr ist die Variation des Lenkgefühls mithilfe verschiedener Varianten zu realisieren. Nur so können Umstiege zwischen den Fahrzeugen sowie eine Beeinflussung der Probanden durch die unterschiedlichen Fahrzeuge vermieden werden.

- OP _3: Variation des Lenkgefühls erfolgt durch verschiedene Varianten

Alles in allem ermöglicht die Nutzung verschiedener Varianten, dass die Probanden blind sind gegenüber den eingestellten Parametrierungen. Dies wird beispielsweise auch von Kudritzki (2000) gefordert. Hierdurch können bspw. wiederholte Bewertungen besonders leicht umgesetzt werden.

- OP _4: Fahrer ist blind gegenüber den zugrundeliegenden Varianten

Tabelle 29. Zusammenstellung der Vor- und Nachteile sowie die mögliche Art der Bewertung in Abhängigkeit der Versuchsumgebung und der Variation des Lenkgefühls.

Variation des Lenkgefühls	Versuchsumgebung		
	Öffentlicher Straßenverkehr	Testgelände	Fahrsimulator
Verschiedene Fahrzeuge	⊕ Einfach umsetzbar ⊖ Beeinflussung durch Gesamtfahrzeug und Varianz in Versuchsumgebung ➤ Art der Bewertung: absolut, falls möglich relativ	⊕ Lenkgefühl als Teil des Gesamtfahrzeugverhaltens standardisiert evaluierbar ⊖ Kein direkter Vergleich möglich ➤ Art der Bewertung: relativ, falls nötig absolut	⊕ Einfacher Wechsel zwischen simulierten Fahrzeugen ⊖ Lenkrad sowie Fahrerhaus nicht austauschbar ➤ Art der Bewertung: relativ
Verschiedene Varianten	⊕ Realitätsnahe, blinde Erprobung ⊖ Schwer umsetzbar, da Straßenfreigabe nötig ➤ Art der Bewertung: relativ	⊕ Unmittelbarer Vergleich in standardisierten Fahrmanövern ⊖ Evtl. Ergebnisse schwer übertragbar ➤ Art der Bewertung: relativ	⊕ Variation leicht umsetzbar ⊖ Ergebnisse von Qualität des Fahrsimulators abhängig ➤ Art der Bewertung: relativ

Fahrsituationen

Zusätzlich zu den allgemeinen Anforderungen des Versuchskonzepts ergeben sich bei der Evaluation auf der Teststrecke weitere Anforderungen:

- FS_5: Anzahl der Fahrmanöver so gering wie möglich

Die Minimierung der Anzahl an Fahrmanövern ist aus mehreren Gründen sinnvoll. Zum einen kann so die Zeit, die die Fahrer zum Vertrautmachen mit den Manövern benötigen, reduziert werden. Zum anderen können sie sich mit wachsender Vertrautheit besser auf die Beurteilung konzentrieren.

- FS_6: Manöver sollten Vorgehen von Experten ähneln

Dies bedeutet zumeist die Nutzung von ISO-Manövern. Alle Manöver müssen jedoch an das fahrerische Können der Normalfahrer angepasst werden. Dies beinhaltet u. a. eine Anpassung der auftretenden Querbeschleunigungen.

- FS_7: Umschalten zwischen Varianten erfolgt innerhalb des Manövers

Das Umschalten zwischen Referenz und Variante sollte innerhalb des Manövers während der Fahrt und ohne Stopp möglich sein (Barthenheier, 2004). Nur so kann die Vergleichbarkeit maximiert werden.

Stichprobe

Aufgrund der Variation des Lenkgefühls durch die Nutzung verschiedener Varianten können die aus Studie I abgeleiteten Anforderungen (SP_4, SP_5) eliminiert werden. Schließlich ist bei der Erprobung innerhalb eines Fahrzeugs keine Beeinflussung durch unterschiedliche Einstellungen zu den Versuchsfahrzeugen möglich.

Bewertungskriterien

Im ersten Schritt wurden die vier NVH-Kriterien *flattern*, *haken*, *knarzen* und *vibrieren* ausgeschlossen. Einerseits zeigte sich in der Faktorenanalyse, dass der Ausschluss von *flattern* und *knarzen* zu einer Erhöhung von Cronbachs α führen würde. Andererseits werden diese Aspekte mit dem Einzug der hybriden und letztendlich der elektrischen Lenksysteme immer mehr an Bedeutung verlieren. Ferner werden sie maßgeblich durch die Mechanik und Hydraulik des Lenksystems vorgegeben. Bei Relativ-Bewertungen sollte die Variation des Lenkgefühls zur Vereinfachung des Versuchsablaufs jedoch softwareseitig erfolgen, sodass diese Kriterien vernachlässigbar wären. Außerdem wurden die Sub-Kriterien zur *Lenkaktivität* und zum *Rückstellverhalten* im Sinne der Versuchsökonomie wieder zusammengefasst. Bei den beiden Kriterien des Rückstellverhaltens betrug der Zusammenhang zwischen den Gefallen-Bewertungen $r = .878$. Auch beim Korrekturbedarf lagen ähnlich große Zusammenhänge zwischen Lenkwinkel und Häufigkeit vor: r (Kurve) = .884, r (Geradeausfahrt) = .934, r (Querneigung) = .894.

Mit einer erneuten explorativen Faktorenanalyse wurden die Item-Ausschlüsse abgesichert. Hierzu wurde eine Hauptkomponentenanalyse mit den gemittelten, z -transformierten Gefallen-Bewertungen der finalen 19 Kriterien berechnet. Bei den zusammengefassten Kriterien wurden die Bewertungen von *Rückstellverhalten-Lenkswinkel*

und *Korrekturbedarf-Lenkwinkel (Kurve)* beibehalten. Der KMO-Wert von .826 war als „lobenswert“ einzustufen (Hutcheson & Sofroniou, 2009). Die individuellen KMO-Werte waren jedoch erneut kleiner als .5. Die mittlere Korrelation zwischen den Kriterien betrug $r = .433$, was marginal größer war als der ursprüngliche Wert. Die Determinante lag mit $2.217\text{E-}9$ unter der Grenze von .00001 (Field, 2013), weshalb Multikollinearität nicht ausgeschlossen werden konnte. Der Ausschluss einzelner Items wurde nahegelegt. Bartlett's Test auf Sphärizität war signifikant mit $p < .001$.

Die ursprüngliche Extraktion zeigte, dass vier Faktoren Eigenwerte größer als 1 hatten und insgesamt 75,81 % der Varianz erklärten. Entsprechend der Parallel-Analyse (O'Connor, 2000) sollten zwei Faktoren mit zusammen 63,20 % Varianz beibehalten werden. Aus dem Ausschluss der Kriterien folgt im Vergleich somit keine Verringerung der erklärten Varianz. Der erste aus der Promax-Rotation resultierende Faktor *Fahrzeugreaktion und Fahrzeugrückmeldung* erklärte mit seinen zwölf Items 47,68 % der Varianz. Die mittlere Korrelation betrug $r = .586$. Cronbachs α als Maß für die interne Konsistenz betrug .944. Der zweite Faktor *Lenkkraft* erklärte mit seinen acht Kriterien zusätzlich 15,52 % der Varianz und wies ein Cronbachs α von .867 auf. Die mittlere Korrelation betrug $r = .482$. In Anlehnung an die bei Cortina (1993) berichteten Werte, die Cronbachs α bei bestimmten Kombinationen von Itemanzahl und mittlerer Korrelation annehmen kann, war bei beiden Faktoren von Eindimensionalität auszugehen. Auch die Trennschärfen größer als .45 sprachen für einen ausreichenden Zusammenhang zwischen allen Items und den Gesamtwerten der beiden Sub-Skalen. Ein Ausreißer nach unten war *gleichmäßig* mit einer Trennschärfe knapp unterhalb .4.

Auf Basis dieser Ergebnisse konnte der reduzierte Kriterienkatalog zur Erfassung des Lenkgefühls am schweren Nutzfahrzeug herangezogen werden. Anschließend wurde eine Neutralisierung der Bewertungskriterien und der Erklärungen vorgenommen. Auch die Erklärungen wurden so formuliert, dass sie spezifisch auf vom Fahrer wahrnehmbare Aspekte hinweisen. Bisher war anzugeben, inwiefern ein Fahrzeug eine Eigenschaft zeigt. Diese Formulierung könnte die Zustimmungstendenz der Probanden verstärkt haben. Nun sollen die Fahrer angeben, wie sich die Lenkung eines Fahrzeugs hinsichtlich einer Eigenschaft verhält. Um die Veränderung zu erläutern, wird das Kriterium *leichtgängig* mit folgender Erklärung herangezogen: „Um zu lenken ist nur ein geringer Kraftaufwand notwendig.“ Nach der Überarbeitung hieß das Kriterium *Lenkmoment* und hatte die Erklärung: „Achten Sie darauf, wie viel Kraft Sie beim Lenken benötigen.“ Die Bewertungskriterien inkl. Erklärungen und korrespondierenden Fahrsituationen sind in Tabelle 30 zu finden.

Tabelle 30. Bewertungskriterien mit dazugehöriger Erklärung und Fahrsituation.

Nr.	Bewertungskriterium	Erklärung	Spezifische Fahrsituation
1	Agilität	Achten Sie darauf, inwiefern die Fahrzeugreaktion sofort nach Ihren Lenkeingaben eintritt.	Mittlere Lenkwinkel in beide Richtungen bei mittleren bis hohen Geschwindigkeiten
2	Anlenkmoment	Achten Sie darauf, wie viel Kraft Sie zum Anlenken aus der Mitte benötigen.	Kleine Lenkwinkel bei hohen Geschwindigkeiten
3	Direktheit	Achten Sie auf die Größe des Lenkwinkels, den Sie für das Erreichen einer bestimmten Fahrzeugreaktion benötigen.	Alle Lenkwinkelbereiche in allen Geschwindigkeitsbereichen

4	Fahrbahnkontakt	Achten Sie darauf, wie gut Sie sich über die Lenkung mit der Fahrbahn verbunden fühlen. (Negativbeispiel: Stapler, Simulator, Glätte, Nässe)	Mittlere und hohe Geschwindigkeiten
5	Haltemoment	Achten Sie darauf, wie viel Kraft Sie zum Halten des Lenkrads in Kurven benötigen. Der Fokus liegt auf dem eingelenkten Zustand, wenn Sie das Lenkrad nur noch festhalten.	Konstanter Lenkwinkel
6	Hat Spiel	Achten Sie darauf, wie groß der tote Bereich um die Mitte ist. Orientieren Sie sich daran, inwiefern die von Ihnen eingegebenen Lenkwinkel zu einer spürbaren Fahrzeugreaktion führen.	Geradeausfahrt bei hohen Geschwindigkeiten
7	Komfortabel	Achten Sie darauf, wie sehr die Lenkung Ihnen ein komfortables Gefühl vermittelt.	Global
8	Korrekturbedarf	Achten Sie darauf, wie viel Sie korrigieren müssen, um das Fahrzeug auf der Spur zu halten.	Geradeausfahrt / kleine Lenkwinkel bei mittleren Geschwindigkeiten / Querneigung
9	Lenkaktivität (i)	Achten Sie darauf, wie viel Sie beim Lenken kurbeln müssen.	Große Lenkwinkel in beide Richtungen bei niedrigen Geschwindigkeiten
10	Lenkmoment	Achten Sie darauf, wie viel Kraft Sie beim Lenken benötigen.	Große Lenkwinkel bei niedrigen Geschwindigkeiten
11	Lenkmomentverlauf	Achten Sie darauf, inwiefern die Lenkkraft über alle Lenkwinkelbereiche gleichmäßig ansteigt. Hinweis: Der Fokus liegt auf dem Einlenkvorgang aus der Mitte.	Große Lenkwinkel bei niedrigen Geschwindigkeiten
12	Lenkpräzision	Achten Sie darauf, wie genau das Fahrzeug auf Ihre Lenkeingaben reagiert. Orientieren Sie sich daran, wie exakt Sie die vorgegebene Pylonengasse durchfahren.	Engstelle
13	Lenkungsstößigkeit (i)	Achten Sie darauf, wie stark das Lenkrad Störungen der Straße weiterleitet. Orientieren Sie sich daran, wie stark Sie Schläge im Lenkrad spüren können.	Schlechte Fahrbahnoberfläche (Kanaldeckel, ausgeprägte Teerfugen)
14	Mittenzentrierung	Achten Sie darauf, wie deutlich die Geradeausstellung des Lenkrads spürbar ist.	Geradeausfahrt bei mittleren und hohen Geschwindigkeiten
15	Nervös (i)	Achten Sie darauf, wie hektisch und stark das Fahrzeug auf Ihre Lenkeingaben reagiert.	Geradeausfahrt bei hohen Geschwindigkeiten
16	Rückstellverhalten	Achten Sie darauf, wie die Lenkung Sie beim Rücklenken unterstützt. Orientieren Sie sich daran, wie schnell sich das Lenkrad zurück dreht und wie groß der Restlenkwinkel ist.	Nach großen Lenkwinkeln bei niedrigen Geschwindigkeiten
17	Sicher	Achten Sie darauf, wie sehr die Lenkung Ihnen ein sicheres Gefühl vermittelt.	Global
18	Synthetisch (i)	Achten Sie darauf, wie deutlich das Lenkrad Ihnen eine Rückmeldung des Fahrzustandes vermittelt (z. B. Radstellung).	Fahrbahnoberfläche (Fahrbahnmarkierung, Querfuge)
19	Verzögert (i)	Achten Sie darauf, wie viel Zeit zwischen Ihren Lenkbewegungen und der Fahrzeugreaktion liegt.	Mittlere Lenkwinkel in beide Richtungen bei mittleren bis hohen Geschwindigkeiten

* (i): invertiert

Bewertungsskala

Abschließend wurde die Skala angeglichen. Passend zu den umformulierten Erklärungen wurden die beiden Skalen „Niveau“ und „Gefallen“ genannt. Eine genauere Erklärung der beiden Bewertungsschritte soll in der Vorbesprechung nach der Erläuterung der Kriterien erfolgen:

- Niveau: Wie stark nehmen Sie die beschriebene Eigenschaft im Vergleich zur Referenz wahr?
- Gefallen: Wie gut gefällt Ihnen dies im Vergleich zur Referenz?

Wie bereits erläutert, ist für die relative Bewertung eine ungerade Anzahl an Skalenpunkten zu verwenden. Die Skalenmitte bedeutet dann, dass kein Unterschied zwischen den verglichenen Variationen des Lenkgefühls wahrgenommen wurde. Die Skalenpunkte im negativen Bereich stehen für eine schwächere Ausprägung der Eigenschaft bei der aktuellen Variante im Vergleich zur Referenz, die Skalenpunkte im positiven Bereich entsprechend für eine stärkere Ausprägung. Diese Systematik ist entsprechend auf die Gefallen-Bewertungen übertragbar (Abbildung 26).

Niveau						
-3	-2	-1	0	+1	+2	+3
sehr viel weniger	viel weniger	etwas weniger	gleich	etwas mehr	viel mehr	sehr viel mehr

Gefallen						
-3	-2	-1	0	+1	+2	+3
sehr viel schlechter	viel schlechter	etwas schlechter	gleich	etwas besser	viel besser	sehr viel besser

Abbildung 26. Benannte 7-stufige Skalen zur relativen Niveau- und Gefallen-Bewertung.

3.4. Versuchskonzept

3.4.1. Fragebogen

Der Fragebogen zur subjektiven Bewertung des Lenkgefühls schwerer Nutzfahrzeuge durch Normalfahrer besteht aus 19 Bewertungskriterien mit den dazugehörigen Erklärungen (Tabelle 31). Weiterhin werden die Fahrsituationen, in denen die einzelnen Bewertungskriterien erlebt werden können, spezifiziert.

Tabelle 31. Bewertungskriterien mit dazugehöriger Erklärung und Fahrsituation.

Nr.	Bewertungskriterium	Erklärung	Spezifische Fahrsituation
1	Agilität	Achten Sie darauf, inwiefern die Fahrzeugreaktion sofort nach Ihren Lenkeingaben eintritt.	Mittlere Lenkwinkel in beide Richtungen bei mittleren bis hohen Geschwindigkeiten
2	Anlenkmoment	Achten Sie darauf, wie viel Kraft Sie zum Anlenken aus der Mitte benötigen.	Kleine Lenkwinkel bei hohen Geschwindigkeiten
3	Direktheit	Achten Sie auf die Größe des Lenkwinkels, den Sie für das Erreichen einer bestimmten Fahrzeugreaktion benötigen.	Alle Lenkwinkelbereiche in allen Geschwindigkeitsbereichen
4	Fahrbahnkontakt	Achten Sie darauf, wie gut Sie sich über die Lenkung mit der Fahrbahn verbunden fühlen. (Negativbeispiel: Stapler, Simulator, Glätte, Nässe)	Mittlere und hohe Geschwindigkeiten
5	Haltemoment	Achten Sie darauf, wie viel Kraft Sie zum Halten des Lenkrads in Kurven benötigen. Der Fokus liegt auf dem eingelenkten Zustand, wenn Sie das Lenkrad nur noch festhalten.	Konstanter Lenkwinkel
6	Hat Spiel	Achten Sie darauf, wie groß der tote Bereich um die Mitte ist. Orientieren Sie sich daran, inwiefern die von Ihnen eingegebenen Lenkradwinkel zu einer spürbaren Fahrzeugreaktion führen.	Geradeausfahrt bei hohen Geschwindigkeiten
7	Komfortabel	Achten Sie darauf, wie sehr die Lenkung Ihnen ein komfortables Gefühl vermittelt.	Global
8	Korrekturbedarf	Achten Sie darauf, wie viel Sie korrigieren müssen, um das Fahrzeug auf der Spur zu halten.	Geradeausfahrt / kleine Lenkwinkel bei mittleren Geschwindigkeiten / Querneigung
9	Lenkaktivität (i)	Achten Sie darauf, wie viel Sie beim Lenken kurbeln müssen.	Große Lenkwinkel in beide Richtungen bei niedrigen Geschwindigkeiten
10	Lenkmoment	Achten Sie darauf, wie viel Kraft Sie beim Lenken benötigen.	Große Lenkwinkel bei niedrigen Geschwindigkeiten
11	Lenkmomentverlauf	Achten Sie darauf, inwiefern die Lenkkraft über alle Lenkwinkelbereiche gleichmäßig ansteigt. Hinweis: Der Fokus liegt auf dem Einlenkvorgang aus der Mitte.	Große Lenkwinkel bei niedrigen Geschwindigkeiten
12	Lenkpräzision	Achten Sie darauf, wie genau das Fahrzeug auf Ihre Lenkeingaben reagiert. Orientieren Sie sich daran, wie exakt Sie die vorgegebene Pylonengasse durchfahren.	Engstelle
13	Lenkungsstößigkeit (i)	Achten Sie darauf, wie stark das Lenkrad Störungen der Straße weiterleitet. Orientieren Sie sich daran, wie stark Sie Schläge im Lenkrad spüren können.	Schlechte Fahrbahnoberfläche (Kanaldeckel, ausgeprägte Teerfugen)
14	Mittenzentrierung	Achten Sie darauf, wie deutlich die Geradeausstellung des Lenkrads spürbar ist.	Geradeausfahrt bei mittleren und hohen Geschwindigkeiten
15	Nervös (i)	Achten Sie darauf, wie hektisch und stark das Fahrzeug auf Ihre Lenkeingaben reagiert.	Geradeausfahrt bei hohen Geschwindigkeiten
16	Rückstellverhalten	Achten Sie darauf, wie die Lenkung Sie beim Rücklenken unterstützt. Orientieren Sie sich daran, wie schnell sich das Lenkrad zurück dreht und wie groß der Restlenkwinkel ist.	Nach großen Lenkwinkeln bei niedrigen Geschwindigkeiten

17	Sicher	Achten Sie darauf, wie sehr die Lenkung Ihnen ein sicheres Gefühl vermittelt.	Global
18	Synthetisch (i)	Achten Sie darauf, wie deutlich das Lenkrad Ihnen eine Rückmeldung des Fahrzustandes vermittelt (z. B. Radstellung).	Fahrbahnoberfläche (Fahrbahnmarkierung, Querfuge)
19	Verzögert (i)	Achten Sie darauf, wie viel Zeit zwischen Ihren Lenkbewegungen und der Fahrzeugreaktion liegt.	Mittlere Lenkwinkel in beide Richtungen bei mittleren bis hohen Geschwindigkeiten

* (i): invertiert

Die Bewertungskriterien werden um eine Bewertungsskala ergänzt, anhand welcher eine Einstufung der subjektiven Einschätzung vorgenommen wird. Die Skalen zur absoluten und zur relativen Bewertung sind in Abbildung 27 zu finden.

Niveau						
-3	-2	-1	0	+1	+2	+3
sehr viel weniger	viel weniger	etwas weniger	gleich	etwas mehr	viel mehr	sehr viel mehr
Gefallen						
-3	-2	-1	0	+1	+2	+3
sehr viel schlechter	viel schlechter	etwas schlechter	gleich	etwas besser	viel besser	sehr viel besser

Abbildung 27. Benannte 7-stufige Skalen zur relativen Niveau- und Gefallen-Bewertung.

3.4.2. Gesammelte Anforderungen an den Versuchsaufbau

Das Versuchskonzept beinhaltet neben dem Fragebogen auch Anforderungen an den Versuchsaufbau. Diese Anforderungen sind als Checkliste zu verstehen, welche bei der Konzeption eines Fahrversuchs herangezogen werden kann (Tabelle 32). Die vorliegenden Anforderungen gliedern sich in die Bereiche Fahrsituation, Bewertungsvorgehen, Versuchsdesign, Operationalisierung und Stichprobe.

Tabelle 32. Checkliste zum Versuchsaufbau.

Bereich	Anforderungen
Fahrsituationen (FS)	<input type="checkbox"/> FS_1: Gesamter Fahrbetrieb abgebildet <input type="checkbox"/> FS_2: Fahrsituationen entsprechen fahrerischem Können der Normalfahrer <input type="checkbox"/> FS_3: Fahrzeit möglichst kurz <input type="checkbox"/> FS_4: Zu Bewertungskriterien zugehörige Fahrsituationen spezifiziert <input type="checkbox"/> FS_5: Anzahl der Fahrmanöver so gering wie möglich <input type="checkbox"/> FS_6: Manöver sollten Vorgehen von Experten ähneln <input type="checkbox"/> FS_7: Umschalten zwischen Varianten erfolgt innerhalb des Manövers
Bewertungsvorgehen (BV)	<input type="checkbox"/> BV_1: Vorstellung aller Bewertungskriterien vor der Fahrt <input type="checkbox"/> BV_2: Aufforderung an Proband, Konzentration auf Lenkung zu richten <input type="checkbox"/> BV_3: Aufforderung an Proband, konstantes Fahrverhalten aufzuweisen <input type="checkbox"/> BV_4: Erklären des Kriteriums vor korrespondierender Fahrsituation <input type="checkbox"/> BV_5: Begründete Beurteilung anhand Skala während oder nach Fahrsituation <input type="checkbox"/> BV_6: Versuchsleiter wiederholt Bewertung <input type="checkbox"/> BV_7: Pro Kriterium mindestens Beantwortungszeitraum von 60 sec verfügbar
Versuchsdesign (VD)	<input type="checkbox"/> VD_1: Versuchsdesign mit Messwiederholung <input type="checkbox"/> VD_2: Wiederholte Bewertung von Fahrzeugen bzw. Varianten
Operationalisierung (OP)	<input type="checkbox"/> OP_1: Bewertung erfolgt als relativer Vergleich unter Nutzung einer Referenz <input type="checkbox"/> OP_2: Teststrecke als Versuchsumgebung <input type="checkbox"/> OP_3: Variation des Lenkgefühls erfolgt durch verschiedene Varianten <input type="checkbox"/> OP_4: Fahrer ist blind gegenüber den zugrundeliegenden Varianten
Stichprobe (SP)	<input type="checkbox"/> SP_1: Stichprobengröße mathematisch bestimmt <input type="checkbox"/> SP_2: Teilnahmevoraussetzungen: 3 Jahre Führerscheinbesitz, jährliche Fahrleistung von 30.000 km, haupt- oder nebenberuflicher Lkw-Fahrer, Erfahrung mit Lkw-Typ, ausreichende Sprachkenntnisse <input type="checkbox"/> SP_3: Demographische Variablen kontrolliert

4. Anwendung des Versuchskonzepts: Bewertung des Lenkgefühls verschiedener Varianten der Lenkunterstützung (Studie III)

4.1. Ziele & Hypothesen

Wie im vorherigen Kapitel vorgestellt, wurde ein Versuchskonzept zur subjektiven Bewertung des Lenkgefühls am schweren Nutzfahrzeug durch Normalfahrer entwickelt. Das Versuchskonzept soll nun angewendet werden, um Erkenntnisse über das Lenkgefühl zu gewinnen. Diese sollen im nächsten Schritt dazu dienen, Gestaltungsmöglichkeiten abzuleiten.

Das Ausmaß, wie stark das Lenkgefühl an die Fahrerbedürfnisse angepasst werden kann, wird maßgeblich durch die Möglichkeit zur freien Gestaltung der Lenkunterstützung bestimmt. Eine variable Lenkunterstützung bedeutet beispielsweise, dass die Höhe des Unterstützungsmoments an die Fahrgeschwindigkeit angepasst werden kann. Nur so können niedrige vom Fahrer aufzuwendende Lenkmomente bei langsamer Fahrt und höhere Lenkmomente bei schnellerer Fahrt realisiert werden (Anand, 2014; Bertolini & Hogan, 1999). Ferner besteht die Möglichkeit, den Charakter einer Lenkung durch eine entsprechende Parametrierung frei zu gestalten (Henrichfreise, Jusseit & Niessen, 2003). Dabei hat das Verhältnis zwischen Lenkmoment und Lenkwinkel einen besonders großen Einfluss auf das Lenkgefühl (Newberry, Griffin & Dowson, 2007). Weiterhin vermittelt das Lenkmoment auch Informationen über den aktuellen Fahrzustand, was insbesondere im fahrdynamischen Grenzbereich von Bedeutung ist (Neukum, Paulig, Frömmig & Henze, 2009).

Eine variable Lenkunterstützung kann in schweren Nutzfahrzeugen heute nur mithilfe von hybriden Lenksystemen realisiert werden. Ein mit einem solchen System ausgestattetes Fahrzeug wurde bereits im Rahmen von Studie II evaluiert (Kapitel 3.3.2). Es zeigte sich unter anderem, dass die extreme Leichtgängigkeit beim Rangieren stark polarisierte. Da bisher kaum gesicherte Ergebnisse zur Beurteilung verschiedener Lenkmomentniveaus im Lkw vorliegen, sind sehr umfassende Erkenntnisse nötig (siehe Kapitel 2.4). Deshalb lag der Fokus der Arbeit auf der Variation und Evaluation des vom Fahrer aufzuwendenden Lenkmoments, um das Potential hybrider Lenksysteme zukünftig optimal nutzen zu können. Insbesondere sollen Rückschlüsse von der Parametrierung der Lenkunterstützung auf das subjektiv empfundene Lenkgefühl gezogen werden können. Im Gegensatz zu den vorherigen Studien wurden verschiedene Parametrierungen des Lenksystems in einem einzigen Fahrzeug abgebildet. Die Erprobung fand auf einem Testgelände anhand synthetischer Fahrmanöver statt.

Bevor jedoch aus den abgegebenen Bewertungen Erkenntnisse zum Lenkgefühl abgeleitet werden, soll das Versuchskonzept abschließend untersucht werden. Mit diesem war von einer zufriedenstellenden Objektivität, Reliabilität und Validität der Messung auszugehen. Eine entsprechende Überprüfung stand im Fokus der ersten drei Hypothesen. Zuerst sollte der verwendete Fragebogen genauer untersucht werden. Hierzu wurde wie in den vorherigen Studien eine explorative Faktorenanalyse sowie eine Item- und Reliabilitätsanalyse durchgeführt. Für Hypothese 1 galt:

H1: Der Fragebogen weist eine ausreichende psychometrische Qualität auf.

Ferner wurde untersucht, wie die Niveau- und Gefallen-Bewertungen zusammenhängen. Die Niveau-Bewertung ist eine quasi-objektive Einschätzung der tatsächlichen Gegebenheiten (Decker, 2009). Sie wird im nächsten

Schritt in eine individuelle Gefallen-Bewertung übersetzt. Diese stellt das eigentliche Lenkgefühl als subjektive, komplexe Erfahrung dar (Harrer & Pfeffer, 2013). Auch wenn persönliche Vorlieben individuell sind, liegen dem Mechanismus dennoch allgemeingültige Muster zugrunde. So deckte Zschocke (2009) lineare und quadratische Zusammenhänge zwischen den Niveau- und Gefallen-Bewertungen der Variablen auf. Diese wiederum sind die Basis für Gestaltungsempfehlungen. Für den folgenden Versuch wurden ebenfalls Abhängigkeiten der Gefallen- von den Niveau-Bewertungen angenommen, weshalb für Hypothese 2 galt:

H2: Die Gefallen-Bewertungen lassen sich in Abhängigkeit von den Niveau-Bewertungen vorhersagen.

Abschließend sollte die Zuverlässigkeit der Bewertungen beleuchtet werden. In Studie II wurde bereits die Konstanz der Bewertungen über die Dauer eines Versuchs gezeigt. Dieses Ergebnis sollte auch für die Bewertung verschiedener Varianten repliziert werden. Bei einer Variation des Lenkgefühls über Software-Änderungen sind die Probanden blind gegenüber den vom Versuchsleiter vorgenommenen Einstellungen. So kann, im Gegensatz zu den vorherigen beiden Studien, eine wiederholte Bewertung von Varianten ohne das Wissen der Probanden erfolgen (Barthenheier, 2004). Sind die Einschätzungen der Probanden zuverlässig, zeigen sich keine signifikanten Unterschiede zwischen ihnen. Hieraus folgte für Hypothese 3:

H3: Die wiederholte Bewertung der gleichen Variante der Lenkunterstützung führt nicht zu unterschiedlichen Bewertungen.

Nach diesen methodischen Fragestellungen sollten konkrete Erkenntnisse bezüglich der Wahrnehmung und Bewertung der Varianten gewonnen werden. Hierzu wurde untersucht, ob die zur Bewertung herangezogenen Varianten der Lenkunterstützung von den Fahrern differenziert werden konnten. Die Varianten wurden auf Basis der vorliegenden Erkenntnisse zu Unterschiedsschwellen (Buschardt, 2003; Hiraoka, Hioki & Kumamoto, 2008; Newberry et al., 2007; Wolf, 2009) so gestaltet, dass Unterschiede wahrnehmbar sein sollten. Hieraus wiederum folgen unterschiedliche Niveau-Bewertungen. Deshalb lautete Hypothese 4 wie folgt:

H4: Die objektiv vorhandenen Unterschiede in der Höhe des aufzuwendenden Lenkmoments spiegeln sich in entsprechend unterschiedlichen Niveau-Bewertungen wider.

Anschließend sollten die Bewertungen der Varianten zur Gewinnung von Auslegungsempfehlungen dienen. Hierzu wurde die Abhängigkeit der subjektiven Bewertungen von den objektiven Parametern untersucht. An dieser Stelle ist jedoch zwischen direkt und indirekt vom Unterstützungsmoment abhängenden Kriterien zu unterscheiden (Kapitel 4.2.5). Bei direkt abhängigen Bewertungskriterien, z. B. *leichtgängig*, ist das vom Fahrer aufzuwendende Lenkmoment ein sehr starker Einflussfaktor (Rothhämel et al., 2011b). Bei indirekt abhängigen Bewertungskriterien, z. B. Lenkpräzision, beeinflusst jedoch eine Vielzahl von Kennwerten die Bewertung. Gemäß Harrer (2007) wird die Genauigkeit von sieben verschiedenen Parametern beeinflusst. Riedel und Arbinger (1997) und Rothhämel et al. (2011a) arbeiteten sogar mit einer noch größeren Anzahl und Vielfalt an Kennwerten. Diese weiteren Parameter sollten jedoch im Rahmen dieses Versuchs nicht variiert werden. Deshalb wurden die nachfolgenden Analysen nur für die Kriterien durchgeführt, welche maßgeblich vom Unterstützungsmoment abhängen.

Im ersten Schritt sollte der Zusammenhang zwischen den Niveau-Bewertungen und den Parametern des Lenksystems, welche den Varianten zugrunde liegen, untersucht werden. Es war anzunehmen, dass die

subjektiven Niveau-Bewertungen von den objektiven Parametern beeinflusst werden (Harrer, 2007; Kim & Yoon, 2015; Zschocke, 2009). Daraus folgte für Hypothese 5:

H5: Die Niveau-Bewertungen der direkt vom aufzuwendenden Lenkmoment abhängigen Bewertungskriterien lassen sich in Abhängigkeit von Parametern der Lenkungsauslegung vorhersagen.

Abschließend wurde auch die Abhängigkeit der Gefallen-Bewertung von den objektiven Kennwerten überprüft. Diese Ergebnisse lassen einerseits unmittelbare Ableitungen von Gestaltungsempfehlungen zu. Andererseits ermöglichen sie auch Rückschlüsse zur Verschlankung der Befragungsmethode, wenn der Fokus einer Untersuchung auf der Aufdeckung von Optimierungspotential liegt. Schließlich wäre das Erfassen von Niveau-Bewertungen überflüssig, wenn starke Zusammenhänge mit den Gefallen-Bewertungen vorlägen. Hieraus ergibt sich für Hypothese 6:

H6: Die Gefallen-Bewertungen der direkt vom aufzuwendenden Lenkmoment abhängigen Bewertungskriterien lassen sich in Abhängigkeit von Parametern der Lenkungsauslegung vorhersagen.

Die Vorbereitung, Durchführung und Auswertung der Studie wurde von Striegel (2016) im Rahmen ihrer Bachelorthesis unterstützt.

4.2. Methodik

4.2.1. Versuchsdesign

Im Rahmen dieses Versuchs sollten verschiedene Varianten der Lenkunterstützung evaluiert werden. Zuerst musste eine geeignete Anzahl an zu vergleichenden Varianten festgesetzt werden. Hierbei war zu beachten, dass das gesamte Spektrum möglicher Auslegungen abgedeckt wird. Außerdem war sicher zu stellen, dass sich die Probanden nicht auf einzelne Varianten als Favoriten festlegen können (Gies & Marusic, 2000). Nichtsdestotrotz mussten Aspekte der Versuchsökonomie beachtet werden. Hieraus resultierten für den Faktor *Variante* 5 Stufen (V1 – V5) (within-subject). Zur Überprüfung der Stabilität der Urteile wurde pro Kriterium eine Variante doppelt beurteilt. Der hierfür eingeführte zweite Faktor *Wiederholung* war dreistufig (within-subject). Handelte es sich um die erstmalige Beurteilung, lag keine Wiederholung vor (nein). Die beiden wiederholten Varianten wurden entweder direkt hintereinander evaluiert (unmittelbar) oder als erste und letzte Variante innerhalb eines Kriteriums (getrennt). Die erzielten Stichprobengrößen der Wiederholungen ergaben sich daraus, dass pro Bewertungskriterium jeweils nur eine Variante wiederholt wurde.

Somit lag für diesen Fahrversuch ein zweifaktorielles within-subject Design vor (Tabelle 33). Die gemessenen abhängigen Variablen waren die subjektiven Beurteilungen der Probanden.

Tabelle 33. Versuchsdesign unter Angabe der Stichprobengrößen.

Faktor 2: Wiederholung	Faktor 1: Variante				
	V1	V2	V3	V4	V5
Nein	$n = 600$	$n = 600$	$n = 600$	$n = 600$	$n = 600$
Unmittelbar	$n = 75$	$n = 75$	$n = 39$	$n = 38$	$n = 75$
Getrennt	$n = 75$	$n = 56$	$n = 55$	$n = 55$	$n = 57$

4.2.2. Stichprobe

Die benötigte Stichprobengröße wurde mit G*Power 3.1.9.2 (Faul, 2007; Faul, 2009) festgesetzt. Die zur Berechnung benötigten erwarteten Effektstärken wurden aus Studie II abgeleitet und lagen durchschnittlich im mittleren Bereich. Diese Effektstärken stammten aus einem Vergleich verschiedener Fahrzeuge in realen Fahrsituationen. Somit war nicht klar, inwiefern sie auf den vorliegenden Versuch übertragbar sind. Einerseits könnten sie aufgrund der erschwerten Vergleichbarkeit durch die zeitliche Distanz unterschätzt sein. Andererseits könnten die wahrgenommenen Unterschiede aufgrund der für die Fahrer beobachtbaren Fahrzeugunterschiede allgemein überschätzt sein. Deshalb wurde erneut von mittleren Effektstärken ausgegangen.

Im Gegensatz zu den vorherigen Studien wurde die Teststärke, also die Wahrscheinlichkeit des Aufdeckens vorhandener Unterschiede, höher angesetzt. Deshalb wurde $1-\beta = .95$ gewählt. Der Vergleich der Beurteilungen der fünf verschiedenen Varianten sollte mithilfe von Varianzanalysen mit Messwiederholung erfolgen. Eine angenommene Effektstärke von $f = .25$ setzte eine Stichprobengröße von $N = 31$ voraus. Für die geplanten Regressionen zwischen den subjektiven und objektiven Kennwerten waren bei einem Prädiktor mit $f^2 = .15$ zur Aufdeckung eines mittleren Effekts $N = 89$ Probanden nötig.

Schlussendlich nahmen $N = 75$ männliche Probanden an dem Versuch teil. Sie wurden über Zeitungsanzeigen, Flyer und soziale Medien rekrutiert und erhielten eine monetäre Aufwandsentschädigung. Die Auswahl der Probanden erfolgte anhand der Teilnahmevoraussetzungen (SP_2). Da in diesem Versuch Fahrmanöver absolviert werden mussten, wurde die Dauer des notwendigen Führerscheinbesitzes auf fünf Jahre angehoben. Die Probanden waren im Durchschnitt 45,5 Jahre alt ($SD = 10,6$). Sie waren seit 23 Jahren ($SD = 11,3$) im Besitz ihrer Fahrerlaubnis. Bei einer durchschnittlichen wöchentlichen Fahrzeit von 33,6 h ($SD = 14,1$) absolvierten die Probanden eine jährliche Fahrleistung von 71.000 km ($SD = 52.000$). Die Gesamt-Fahrleistung der Teilnehmer betrug durchschnittlich 1.565.000 km ($SD = 1.317.000$). Einer Tätigkeit als haupt- oder nebenberuflicher Fahrer gingen 70,7 % im Verteilerverkehr und 71,6 % im Fernverkehr nach.

4.2.3. Variation des Lenkgefühls

Gemäß dem Versuchskonzept erfolgte die Variation des Lenkgefühls in dieser Studie mithilfe verschiedener Varianten der Lenkunterstützung (OP_3). Hierzu wurde als Versuchsfahrzeug eine Sattelzugmaschine mit 450 PS

und 18 t zulässigem Gesamtgewicht benutzt. Der angehängte Euro-Sattelaufleger war nicht beladen, um die Beschleunigungsphasen zu verkürzen und sicherheitskritische Situationen bei ungeübten Fahrern oder Nässe zu verhindern.

Um die nötigen Anpassungen der Lenkunterstützung in einem Fahrzeug abbilden zu können, war dieses mit einem hybriden Lenksystem ausgerüstet. Eine zusätzliche Versuchstechnik ermöglichte das Umschalten zwischen den Varianten. Hierzu wurde ein Drehschalter genutzt, welcher an einem für den Fahrer nicht sichtbaren Ort nahe dem Beifahrersitz angebracht war und vom Versuchsleiter betätigt wurde.

Gemäß dem Versuchsdesign, sollten fünf für die Fahrer differenzierbare Varianten der Lenkungsauslegung evaluiert werden. Ferner sollten, wie bei Riedel und Arbinger (1997), neben kleineren auch größere Unterschiede im Lenkverhalten abgebildet werden. Dies war ohnehin nötig, um für das gesamte Geschwindigkeitsspektrum plausible Varianten anbieten zu können. In dem soeben beschriebenen Versuchsfahrzeug konnten die Parameter Lenkmomentniveau, Lenkmomentgradient und aktiver Rücklauf variiert werden. Die Unterschiede, die für die drei Parameter anzusetzen waren, sollten aus der Literatur abgeleitet werden. Hier waren hauptsächlich Unterschiedsschwellen in Bezug auf Lenkmomentenänderungen zu finden.

Dabei wird stets die Frage aufgeworfen, ob die Schwelle zur Wahrnehmung von Änderungen des Lenkmoments abhängig vom Ausgangslenkmoment ist (Barthenheier, 2004; Neukum et al., 2009). Dies legt Webers Gesetz von 1846 (Weber, 2012) über die psychophysikalischen Gesetzmäßigkeiten nahe. Newberry et al. (2007) zeigten, dass die Wahrnehmungsschwellen des Lenkmoments signifikant größer wurden mit größer werdender Ausgangskraft. In 79 % der Fälle wurde ein Unterschied von 15 % im Vergleich zur Ausgangskraft wahrgenommen. Demgegenüber wiesen andere Arbeiten (Buschardt, 2003) keine stringente Abhängigkeit der Differenzschwellen vom Ausgangslenkmoment nach. Alles in allem wurden Lenkmomentänderungen von 0,4 bis 0,5 Nm von den Probanden gerade noch wahrgenommen. Änderungen ab 0,8 Nm wurden zuverlässig detektiert.

Studien zeigten jedoch auch, dass kleine Unterschiede eher wahrgenommen werden, wenn es sich um Lenkmomentenerhöhungen statt um -verringerungen handelte (Wolf, 2009). Eine Verminderung oder Erhöhung des Ausgangslenkmoments von 2 Nm wurde ab 0,4 Nm von 50 % der Probanden entdeckt. Bei einem Ausgangsmoment von 4 Nm wurde eine Erhöhung ab 0,6 Nm von 50 % der Probanden entdeckt, eine Verringerung jedoch erst ab 1,4 Nm. Darüber hinaus fanden Hiraoka et al. (2008), dass bei einem Ausgangslenkmoment von 1,3 Nm eine Verringerung um 0,75 Nm von ca. 80 % der Probanden wahrgenommen wurde.

Insgesamt zeichnet sich die Tendenz ab, dass Wahrnehmungsschwellen abhängig von der Höhe des Ausgangsmoments sind. Da sich die Ergebnisse im Wesentlichen auf Untersuchungen in Mock-ups, Laborprüfständen und Fahrsimulatoren beziehen und insgesamt synthetischer Natur sind, ist die Übertragbarkeit auf reale Fahrsituationen nicht sichergestellt.

Somit wurden die maximalen Lenkmomenten-Niveaus der fünf Varianten wie folgt festgesetzt (Abbildung 28): Die mittlere V3 diente als Referenz (Breuer, 2000; Buschardt, 2003; Decker, 2009) und wies ein maximales vom Fahrer aufzuwendendes Lenkmoment von 3,2 Nm auf. Zwei Varianten wurden so gewählt, dass sie einen

Unterschied von -0,7 Nm (V2) und +0,7 Nm (V4) zur Referenz aufwiesen. Diese Differenz sollte gerade noch wahrnehmbar sein. Die Unterschiede der beiden anderen Varianten zur Referenz waren mit -1,7 Nm (V1) und +1,1 Nm (V5) deutlich größer. So wurde sichergestellt, dass die Probanden in jeden Fall Unterschiede feststellen konnten und das Versuchskonzept validiert werden konnte.

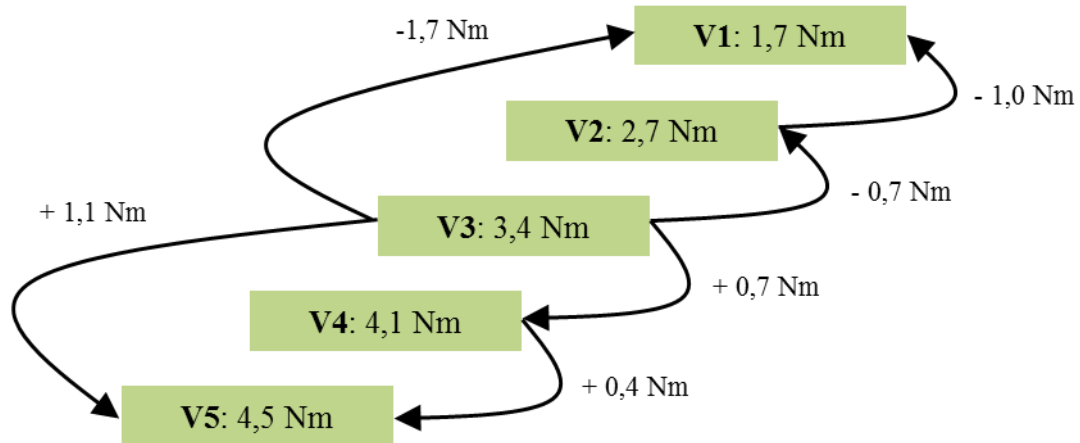


Abbildung 28. Vom Fahrer aufzuwendende Lenkmomente der fünf Varianten.

Mit der Variation der Lenkunterstützung ging auch eine Variation des Rückstellmoments einher. Dieses betrug bei V1 0 Nm, was mit keinem aktiven Rücklauf gleichzusetzen ist. Bei V2 war es mit 4,3 Nm am höchsten. Bei V3 betrug das Rückstellmoment 3,5 Nm, bei V4 3,1 Nm und bei V5 2,5 Nm.

4.2.4. Versuchsumgebung und Festlegen relevanter Fahrsituationen

Um die verschiedenen Varianten möglichst unmittelbar vergleichen zu können, fand der Versuch auf einem Testgelände (OP_2) in synthetischen Fahrmanövern statt. Bei der Auswahl der Fahrsituationen sollte auf ISO-Manöver zurückgegriffen werden. In der Fahrzeugerprobung werden open-loop und closed-loop Manöver unterschieden. In open-loop Tests übernimmt der Fahrer nur die Steuerung des Fahrzeugs. Da die Lenkwinkleingabe des Fahrers unabhängig von der Fahrzeugreaktion ist, können fahrerunabhängige Fahrzeugparameter generiert werden (Kraft, 2011). Diese Tests sind zur Messung objektiver Parameter geeignet. In closed-loop Tests hingegen übernimmt der Fahrer nicht nur eine steuerungs-, sondern auch eine regelungstechnische Aufgabe, indem er auf das Fahrzeugverhalten reagiert. Dies beinhaltet beispielsweise die Korrektur von Ist-Spur auf Soll-Spur. Da das Lenkgefühl aus der Interaktion zwischen Fahrer und Fahrzeug resultiert, findet die subjektive Bewertung im Rahmen von closed-loop Tests statt (Gil Gómez et al., 2015). Die somit zur Verfügung stehenden standardisierten Manöver mussten an das fahrerische Können der Normalfahrer (FS_2) sowie an die örtlichen Gegebenheiten angepasst werden. Zusätzlich war es nötig, neue Manöver zu entwerfen, um alle relevanten Fahrsituationen abbilden zu können (Tabelle 34):

- **Geradeausfahrt:** Besonders im Fernverkehr hat die Geradeausfahrt auf der Autobahn eine große Bedeutung. Sie zeichnet sich durch sehr kleine Lenkwinkel aus. Der Fahrer wird bei der Spurführung durch die selbstständige Zentrierung des Lenkrads unterstützt.

- **Slalom:** Der Slalom war an den Weave-Test (ISO 13674-1:2010) angelehnt. Dieser ist lediglich für Pkw und leichte Nutzfahrzeuge anwendbar. Daher wurde das Manöver durch die Anpassung des Pylonenabstands und eine Verringerung der Fahrgeschwindigkeit für den vorliegenden Versuch adaptiert.
- **Doppelter Spurwechsel:** Das Vorgehen war an den ISO-Spurwechsel (ISO 3888-1:1999) angelehnt. Dieses hochdynamische Manöver wurde jedoch ebenso durch eine Anpassung des Pylonenabstands und eine Verringerung der Fahrgeschwindigkeit entschärft.
- **Oval:** Weitere Fahrsituationen, welche im Verteilerverkehr häufig vorkommen, sind Stadt- und Kreisverkehrsdurchfahrten. Diese sind für den Fahrer aufgrund der Größe der aufzuwendenden Lenkwinkel besonders relevant. In diesen Situationen ist ferner das Rückstellverhalten von großer Bedeutung.
- **Rangieren:** Zu Beginn bzw. Ende eines jeden Transports muss der Fahrer zum Be- und Entladen fahren. Hierbei ist es oft nötig, mit dem Lkw rückwärts-seitwärts an eine Laderampe heranzufahren.

Tabelle 34. Beschreibung der Fahrmanöver unter Angabe der Fahrgeschwindigkeit v , des maximalen Lenkwinkels δ und der maximalen Querbeschleunigung a_y .

Manöver	Beschreibung	Use Case	Vorgehen	v [km/h]	δ [°]	a_y [m/s ²]
Geradeausfahrt	Geradeausfahrt innerhalb Fahrspur	Geradeausfahrt	200 m Referenz → Umschaltzone → 240 m Variante	60	15	0,3
Slalom	Sinusförmige Fahrt um Pylonen (Pylonen-Abstand 42 m)	Kurvige Landstraße	1 Periode Referenz → Umschaltzone → 4 Perioden Variante	45	100	2,7
Doppelter Spurwechsel	Spurwechsel innerhalb Pylonengasse (Abstand zwischen Gassen 42 m, Gassenlänge 25 m)	Spurwechsel, Ausweichen von Hindernissen	1 Zyklus Referenz → Umschaltzone → 2 Zyklen Variante	45	60	1,7
Oval	Folgen einer ovalen Fahrbahnmarkierung (Geradenlänge 25 m, Kurvenradius 12 / 15 m)	Durchfahrt Kreisverkehr	1 Zyklus Referenz → Umschaltzone → 2 Zyklen Variante	15	550	2,2
Rangieren	Rückwärts auf andere Fahrspur zurückstoßen (35 m langer Korridor)	Rangieren	1 Zyklus Referenz → Umschaltzone → 1 Zyklus Variante	max. 5	500	0,3

Aufgrund des Wechsels zwischen den Varianten war eine Umschaltzone bzw. ein Umschaltpunkt nötig (Abbildung 29). Hier waren keine Lenkbewegungen seitens des Fahrers zulässig. Diese Umschaltzone betrug 60 m für die Manöver Geradeausfahrt, Slalom und doppelter Spurwechsel. Beim Oval fand das Umschalten auf der zweiten Geraden statt. Die Fahrgeschwindigkeit blieb in allen Fällen konstant. Beim Rangieren wurde in der Ausgangsposition im Stillstand umgeschaltet.

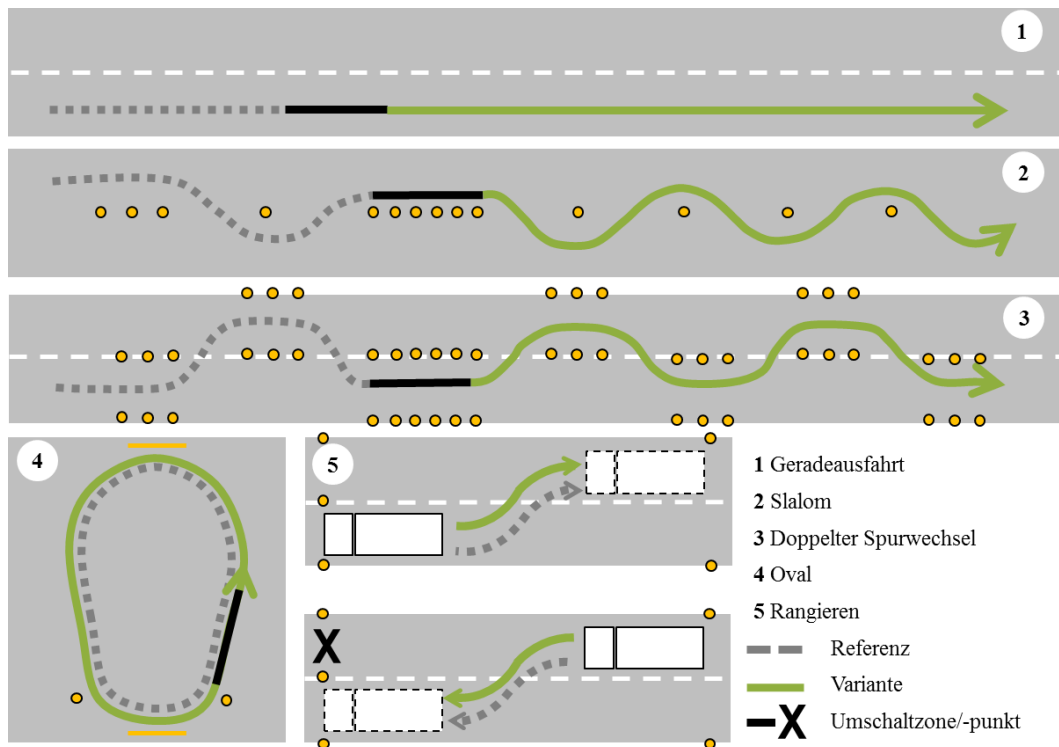


Abbildung 29. Schematischer Ablauf der fünf Fahrmanöver.

Die folgenden Vorkehrungen sollten eine möglichst standardisierte Fahrweise bewirken (Schalz et al., 2002):

- Alle Manöver wurden mit Tempomat absolviert, wodurch die Fahrgeschwindigkeit über alle Probanden hinweg konstant war. Für das Rangieren wurde der sog. Rangiermodus benutzt, welcher eine maximale Geschwindigkeit von 5 km/h zulässt.
- Die Fahrspur wurde, sofern möglich, eng mit Pylonen abgesteckt.
- Die Probanden wurden aufgefordert, möglichst reproduzierbar zu fahren und ihr Fahrverhalten nicht während des Manövers zu ändern.
- Falls das Fahrverhalten der Probanden nicht den Vorgaben entsprach, wurden sie vom Versuchsleiter darauf hingewiesen und der Durchgang ggf. wiederholt.

4.2.5. Art der Bewertung, Bewertungskriterien und Bewertungsskala

Entsprechend des in Kapitel 3.4 dargestellten Fragebogens erfolgte eine relative Niveau- und Gefallen-Bewertung anhand einer 7-stufigen Relativ-Skala. Um den Versuch in einem akzeptablen zeitlichen Rahmen zu halten, konnte lediglich ein Teil der Bewertungskriterien benutzt werden. Der erste Schritt war somit die Auswahl der Bewertungskriterien, die für die Abstimmung des Lenkgefühls einer hybriden Lenkung am relevantesten sind. Dies geschah in einem iterativen Prozess mit der Auswahl geeigneter Fahrsituationen.

Wie oben beschrieben, wurde in diesem Versuch das Maß der Lenkunterstützung variiert, aber keine Veränderungen an der Lenkübersetzung, Dämpfung etc. vorgenommen. Da die starke Variationsfähigkeit der Lenkunterstützung ein wesentlicher Vorteil der hybriden Lenkung ist, sollte in jedem Fahrmanöver ein Aspekt des vom Fahrer aufzubringenden Lenkradmoments erfasst werden. Dies gilt nicht für die Geradeausfahrt, da die

Höhe der aufzuwendenden Lenkmomente bei den geringen aufzuwendenden Lenkwinkeln einen vernachlässigbaren Einfluss hat. Somit wurden die Kriterien *Anlenkmoment*, *Haltemoment* und *Lenkmoment* zur expliziten Bewertung des Lenkmoments genutzt. Zusätzlich wird das *Rückstellverhalten* erhoben, da dieses ebenfalls systematisch variiert werden konnte.

Zusätzlich sollte in einigen Fahrmanövern ein weiteres Kriterium beurteilt werden, welches indirekt mit dem Lenkmoment zusammenhängt. Da sich hier eine Fülle an Kriterien anbot, musste eine Priorisierung vorgenommen werden. In erster Linie sollten die wesentlichen Aspekte des Lenkgefühls abgedeckt werden. Es wurden außerdem solche Aspekte betrachtet, in welchen hybride Lenkungen mehr Auslegungspotential bieten als rein hydraulische Lenksysteme. Des Weiteren wurden Experteneinschätzungen herangezogen, in welchen Kriterien tatsächlich Unterschiede zwischen den Varianten festgestellt werden könnten.

- Ein zentraler Aspekt des Lenkgefühls ist die *Präzision*. Ein hohes Maß an Genauigkeit führt dazu, dass der Fahrer dem Straßenverlauf mit geringem Korrekturbedarf folgt (Harrer & Pfeffer, 2013). Insbesondere bei geringen aufzuwendenden Lenkmomenten fällt es jedoch schwer, den benötigten Lenkwinkel sauber einzustellen, sodass die Präzision herabgesetzt wird (Pfeffer & Scholz, 2010).
- Leichtgängige Lenkungen könnten bei schnellerer Fahrt das Gefühl vermitteln, dass der Kontakt zur Fahrbahn reduziert ist. Dies ist gleichermaßen unkomfortabel wie sicherheitskritisch, da der Fahrer in Grenzsituationen keine Rückmeldung über das Lenkrad erhält (Anand, 2014). Weiterhin kann ein niedriger subjektiver *Fahrbahnkontakt* zu einem „synthetischen“ Lenkgefühl führen, wie es bei Gabelstaplern oder im Fahrsimulator auftritt.
- Der Fahralltag im Fernverkehr besteht im Wesentlichen aus Geradeausfahrten auf der Autobahn, weshalb der Spurhaltung und somit auch der *Mittenzentrierung* eine besondere Bedeutung zukommt (Shyrokau et al., 2015). Durch die selbstständige Zentrierung des Fahrzeugs bzw. Lenkrads können Korrekturbewegungen und Haltekräfte reduziert werden (Heißing & Brandl, 2002). Dies kann durch die Variation des Lenkmomentgradienten im Anlenkbereich beeinflusst werden.

Somit wurden die folgenden Kriterien ausgewählt: *Anlenkmoment*, *Haltemoment*, *Lenkmoment*, *Rückstellverhalten*, *Lenkpräzision*, *Fahrbahnkontakt* und *Mittenzentrierung*. Mit einer erneuten explorativen Faktorenanalyse wurde diese Auswahl abgesichert. Hierzu wurde mit den gemittelten, z-transformierten Gefallen-Bewertungen aus Studie II eine Hauptkomponentenanalyse berechnet. Für das *Rückstellverhalten* wurden die Bewertungen von *Rückstellverhalten-Lenkwinkel* benutzt. Der KMO-Wert von .732 war als „mittelmäßig“ einzustufen (Hutcheson & Sofroniou, 2009). Sechs der individuellen KMO-Werte lagen unter der Akzeptanzgrenze von .5. Die mittlere Korrelation zwischen den Kriterien betrug $r = .375$. Die ebenfalls berechnete Determinante lag mit .027 über der Grenze von .00001 (Field, 2013), weshalb Multikollinearität ausgeschlossen werden konnte. Bartlett's Test auf Sphärizität war signifikant mit $p < .001$.

Die ursprüngliche Extraktion zeigte, dass zwei Faktoren Eigenwerte größer als 1 hatten und insgesamt 71,49 % der Varianz erklärten. Entsprechend der Parallel-Analyse (O'Connor, 2000) sollten beide Faktoren beibehalten werden. Der erste aus der Promax-Rotation resultierende Faktor *Lenkkraft* erklärte mit seinen vier Items 47,65 % der Varianz. Cronbachs α als Maß für die interne Konsistenz betrug $\alpha = .872$. Die mittlere Korrelation zwischen den Kriterien betrug $r = .631$. Der zweite Faktor *Fahrzeugreaktion und Fahrzeugrückmeldung* erklärte mit seinen drei Kriterien zusätzlich 23,83 % der Varianz und wies ein Cronbachs α von .727 auf. Die mittlere

Korrelation zwischen den Kriterien betrug $r = .471$. Für beide Faktoren war somit Eindimensionalität anzunehmen (Cortina, 1993). Auch die Trennschärfen größer als .6 sprachen für einen ausreichenden Zusammenhang zwischen allen Items und den Gesamtwerten der beiden Sub-Skalen (Kallus, 2010). Einzige Ausnahme war hier das Kriterium *Fahrbahnkontakt* mit einer Trennschärfe knapp unter .4. Die Analysen zeigten insgesamt, dass die ausgewählten Kriterien anhand der Daten aus dem letzten Versuch als brauchbar einzustufen waren. Deshalb wurden die Bewertungskriterien zu den Fahrmanövern zugeordnet (Tabelle 35).

Tabelle 35. Bewertungskriterien mit dazugehöriger Erklärung und Fahrsituation.

Faktor	Bewertungskriterium	Erklärung	Fahrmanöver
Lenkkraft	Lenkmoment (Rangieren)	Achten Sie darauf, wie viel Kraft Sie beim Lenken benötigen.	Rangieren
	Lenkmoment (Slalom)	Achten Sie darauf, wie viel Kraft Sie beim Lenken benötigen.	Slalom
	Anlenkmoment	Achten Sie darauf, wie viel Kraft Sie zum Anlenken aus der Mitte benötigen.	Doppelter Spurwechsel
	Haltemoment	Achten Sie darauf, wie viel Kraft Sie zum Halten des Lenkrads in Kurven benötigen. Der Fokus liegt auf dem eingelenkten Zustand, wenn Sie das Lenkrad nur noch festhalten.	Oval
	Rückstellverhalten	Achten Sie darauf, wie die Lenkung Sie beim Rücklenken unterstützt. Orientieren Sie sich daran, wie schnell sich das Lenkrad zurück dreht und wie groß der Restlenkwinkel ist.	Oval
Fahrzeugreaktion und Fahrzeug-rückmeldung	Lenkpräzision	Achten Sie darauf, wie genau das Fahrzeug auf Ihre Lenkeingaben reagiert. Orientieren Sie sich daran, wie exakt Sie die vorgegebene Pylonengasse durchfahren.	Doppelter Spurwechsel
	Fahrbahnkontakt	Achten Sie darauf, wie gut Sie sich über die Lenkung mit der Fahrbahn verbunden fühlen. (Negativbeispiel: Stapler, Simulator, Glätte, Nässe)	Slalom
	Mittenzentrierung	Achten Sie darauf, wie deutlich die Geradeausstellung des Lenkrads spürbar ist.	Geradeausfahrt

4.2.6. Versuchsablauf

Die Reihenfolge der Manöver, der Kriterien und der Varianten wurde randomisiert. Um den Versuchsablauf dennoch überschaubar zu halten, wurden vier Sets generiert. Im ersten Schritt wurde die Reihenfolge der Manöver randomisiert. So sollte sichergestellt werden, dass eine mögliche Ermüdung der Probanden über die gesamte Versuchsdauer keinen systematischen Einfluss auf die Bewertung bestimmter Manöver hat. Außerdem war anzunehmen, dass die Bewertungen des allerersten Kriteriums im Vergleich weniger reliabel sind, da der Proband

die Spanne der abgebildeten Varianten noch nicht kennt. Im zweiten Schritt wurde bei den Manövern mit zwei Kriterien (Oval, Slalom und Doppelter Spurwechsel) deren Reihenfolge randomisiert. So wurde ein möglicher Reihenfolge-Effekt aufgrund der Gewöhnung an das Manöver vermieden. Im dritten Schritt wurde die Abfolge der einzelnen Varianten randomisiert, um ihre Wiedererkennung zu verhindern. Es sollte den Probanden nicht möglich sein, die fünf verschiedenen Varianten sowie die wiederholte Variante als solche zu identifizieren.

Die Versuche dauerten jeweils ca. 3,5 h und fanden auf einem Prüfgelände in Norddeutschland statt. Die Probanden nahmen einzeln an dem Versuch teil und wurden über den gesamten Zeitraum von einem Versuchsleiter begleitet. Nach einer Sicherheitsunterweisung fuhr der Proband im Versuchsfahrzeug auf die Teststrecke, was gleichzeitig der Gewöhnung an das Fahrzeug diente. Auf der Fahrdynamikfläche wurden die einzelnen Fahrmanöver entsprechend des Sets nacheinander absolviert.

Das Vorgehen wird beispielhaft am doppelten Spurwechsel erläutert (Abbildung 30).

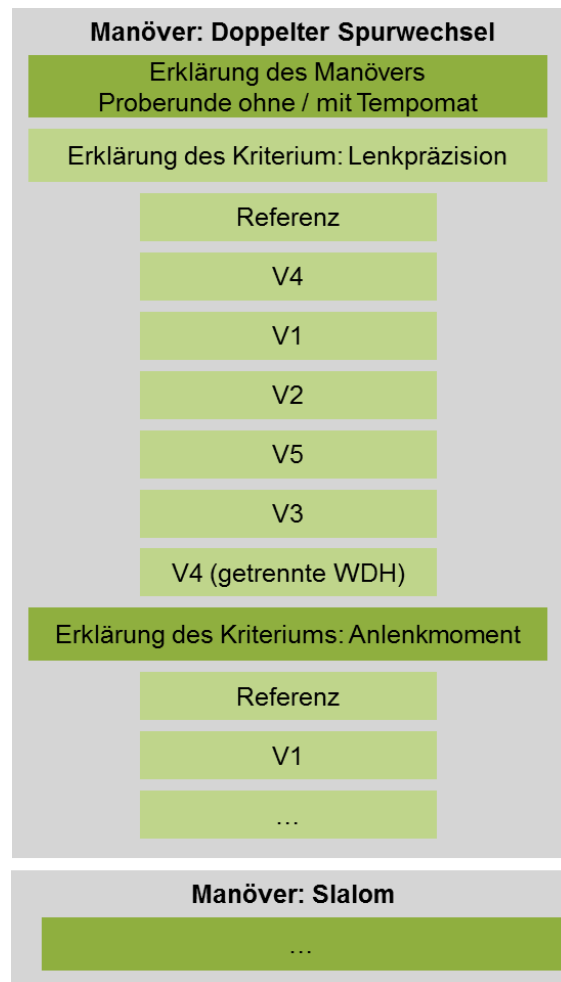


Abbildung 30. Versuchsablauf am Beispiel des doppelten Spurwechsels.

Zu Beginn wurde dem Probanden eine entsprechende Skizze gezeigt und der Use Case erläutert. Jedes Manöver wurde anschließend einmal in Wunschgeschwindigkeit ohne Tempomat und einmal in der Ziel-Geschwindigkeit mit Tempomat durchfahren, damit sich der Proband mit dem Parcours vertraut machen konnte. Im Anschluss

wurde das zu bewertende Kriterium vorgestellt, bspw. *Lenkpräzision*. Der Proband absolvierte den dritten Durchgang komplett in der Referenz und sollte explizit auf die Ausprägung des Kriteriums achten (Chen, 1997; Decker, 2009; Rothhämel et al., 2011a). Nachfolgend wurden alle sechs Varianten (fünf Varianten, eine wiederholte Variante) beurteilt. Hierzu wurde jeweils der erste doppelte Spurwechsel in der Referenz absolviert, um diese stets aufzufrischen. Nach der Umschaltzone und dem Wechsel zur entsprechenden Variante wurden zwei doppelte Spurwechsel absolviert. Die möglicherweise auftretende Veränderung in der Ausprägung und im Gefallen wurde anhand der Skala quantifiziert.

Nach der Beurteilung der *Lenkpräzision* aller sechs Varianten wurde zum nächsten Kriterium übergegangen, welches in diesem Beispiel das *Anlenkmoment* war. Hiernach wurden die restlichen Manöver entsprechend obigem Vorgehen absolviert. Anschließend wurden demographische Daten erhoben.

4.2.7. Auswertungsvorgehen

Die statistische Auswertung erfolgte mit IBM SPSS Statistics 21. Die Daten der Niveau- und Gefallen-Bewertungen wurden zuerst einer deskriptiven Analyse und anschließend einer z -Transformation unterzogen (Kapitel 3.2.2).

Entsprechend der Voraussetzungen wurden die Daten nur bei Stichprobengrößen von $N < 30$ (Bortz & Schuster, 2010) mithilfe des Kolmogorov-Smirnov-Test auf Normalverteilung geprüft. Bei den Varianzanalysen mit Messwiederholung wurde zusätzlich der Mauchly-Test auf Sphärizität durchgeführt. War dieser signifikant, wurde eine entsprechende Korrektur vorgenommen: War $\epsilon < .75$ wurde die Greenhouse-Geisser-Korrektur (Greenhouse & Geisser, 1959) angewandt, bei $\epsilon > .75$ die Huynh-Feldt-Korrektur (Huynh & Feldt, 1976).

Das Signifikanzniveau betrug $\alpha = .05$. Signifikanzwerte unter .05 werden mit * gekennzeichnet, Werte unter .01 mit ** und Werte unter .001 mit ***. Entgegen gängiger Auswertungspraxis wurden die aus den Signifikanztests berechneten p -Werte nach Field (2013) nicht halbiert. Zur Vermeidung der α -Fehler-Kumulierung wurde jeweils eine Benjamini-Hochberg-Korrektur für die Hypothesentests bzw. die explorative Analyse durchgeführt (Benjamini & Hochberg, 1995). Die paarweisen Vergleiche im Rahmen der Ableitung der Gestaltungsempfehlungen wurden separat nach Sidak korrigiert. Für paarweise Vergleiche wurden Effektstärken berechnet. Es gelten gemäß Cohen (1988) für d die Grenzen .2 für einen kleinen, .5 für einen mittleren und .8 für einen großen Effekt. Für r sind die Grenzen .1 für einen kleinen, .3 für einen mittleren und .5 für einen großen Effekt. Für ANOVAs wurde das partielle η^2 als Effektstärke berechnet. Es hat die Grenzen .01 für einen kleinen, .06 für einen mittleren und .14 für einen großen Effekt.

Die psychometrische Untersuchung des Fragebogens erfolgte mithilfe einer explorativen Faktoren-, Item- und Reliabilitätsanalyse. Ferner werden Angaben zur Objektivität, Reliabilität und Validität gemacht.

Bei der Berechnung der Regressionsmodelle wurde wie folgt vorgegangen: Zuerst wurde mithilfe einer Kurvenanpassung untersucht, welche Modelle geeignet sind zur Vorhersage der abhängigen Variable. Aus Gründen der Ökonomie sollten jedoch möglichst einfache Modelle herangezogen werden. Führt mehrere Modelle zu signifikanten Vorhersagen, wurden sie im nächsten Schritt gegeneinander verglichen. Führt die Hinzunahme

einer Komponente zu einem Anstieg der erklärten Varianz, wurde sie beibehalten. Andernfalls wurde das jeweils einfachere Modell angenommen. Ferner wurden folgende Voraussetzungen überprüft (Field, 2013):

- Die Unabhängigkeit angrenzender Residuen wurde mit dem Durbin-Watson-Test abgesichert (Durbin & Watson, 1951). Der berechnete Kennwert kann Werte zwischen 0 und 4 annehmen. In Abhängigkeit der Stichprobengröße und der Anzahl an Prädiktoren wurden die entsprechenden kritischen Werte festgesetzt (Cummins, 2012).
- Anschließend wurden die Ausreißer analysiert. Gemäß Stevens (2002) soll diese Überprüfung nicht dazu führen, die Güte des Modells durch den Ausschluss bestimmter Fälle zu verbessern. Deshalb sollen nur dann Fälle eliminiert werden, wenn die Cook-Distanz > 1 ist (Cook & Weisberg, 1982).

Obige Voraussetzungen waren für keines der berechneten Regressionsmodelle verletzt.

4.3. Ergebnisse

4.3.1. Summative Evaluation des Fragebogens

H1: Der Fragebogen weist eine ausreichende psychometrische Qualität auf.

Faktoren-, Item- und Reliabilitätsanalyse

Da die Zuverlässigkeit der berechneten Faktorenlösung in den vorherigen Fahrversuchen aufgrund der geringen Stichprobe anzuzweifeln war, wird auf diese Problematik als erstes eingegangen.

Es wurde erneut eine Hauptkomponentenanalyse mit den z -transformierten Gefallen-Bewertungen berechnet. Die ursprüngliche Extraktion zeigte, dass zwei Faktoren Eigenwerte größer als 1 hatten. Die aufgeklärte Varianz betrug 56,50 %. Auf Basis der Parallelanalyse nach O'Connor (2000) wurden beide Faktoren beibehalten.

Vor weiteren Analysen wurden die Kommunalitäten berechnet. Diese lagen zwischen .4 und .8. Nach Mundfrom et al. (2009) führen diese bei dem vorliegenden Verhältnis von Variablen zu Faktoren zu einer ausgezeichneten Kongruenz. Auch nach den Empfehlungen von Guadagnoli und Velicer (1988) war die zugrundeliegende Stichprobengröße von $n = 357$ bei den vorliegenden Faktorladungen $> .5$ ausreichend. Somit war von einer Belastbarkeit der Faktorenlösung auszugehen.

Der KMO-Wert von .752 war gemäß Hutcheson und Sofroniou (2009) als „mittelmäßig“ einzustufen. Alle acht individuellen KMO-Werte waren höher als der geforderte Wert von .5 (Field, 2013). Die mittlere Korrelation betrug $r = .258$. Die Determinante lag mit .163 über der Grenze von .00001 (Field, 2013), weshalb Multikollinearität ausgeschlossen werden konnte. Bartlett's Test auf Sphärizität war signifikant mit $p < .001$, was für einen Unterschied zwischen der Korrelationsmatrix der vorliegenden Daten und der Einheitsmatrix steht.

Da eine Abhängigkeit der Faktoren untereinander anzunehmen war, wurde die oblique Promax-Rotation genutzt (Kline, 2000). Es ergab sich eine interpretierbare Verteilung der Items auf die beiden Faktoren (Tabelle 36). Der Faktor *Lennkraft* erklärte mit seinen fünf Items 32,08 % der Varianz. Cronbachs α als Maß für die interne Konsistenz betrug .705. Die mittlere Korrelation lag bei $r = .324$. Der zweite Faktor *Fahrzeugreaktion- und Fahrzeugrückmeldung* erklärte mit seinen drei Items zusätzlich 24,42 % der Varianz und wies ein Cronbachs α

von .596 auf. Die mittlere Korrelation zwischen den Items betrug $r = .331$. Bei Cortina (1993) sind lediglich Hinweise für Skalen mit größeren Itemanzahlen zu finden. Jedoch legen diese Angaben den Schluss nahe, dass auch hier bei beiden Faktoren von Eindimensionalität auszugehen ist.

Tabelle 36. Zweifaktorielle Lösung unter Angabe der internen Konsistenz, Faktorladung und Trennschärfe.

Faktor	Cronbachs α	Bewertungskriterium	Faktorladung	Trennschärfe
Lenkaktivität	.705	Anlenkmoment	.664	.498
		Haltemoment	.761	.526
		Lenkmoment (Rangieren)	.655	.365
		Lenkmoment (Slalom)	.665	.477
		Rückstellverhalten	.640	.445
Fahrzeugreaktion	.596	Fahrbahnkontakt	.782	.468
		Lenkpräzision	.736	.448
		Mittenzentrierung	.581	.307

Die Analyse der Trennschärfen (korrigierte Item-Skala-Korrelation) zeigte für keines der Items einen Wert unter der Grenze von .3 an (Field, 2013). Somit hingen alle Items in einem ausreichenden Maß mit den Gesamtwerten der beiden Faktoren zusammen.

Wie bei den Studien I und II erfolgte auch hier eine Analyse der Häufigkeiten (Tabelle 37). Aufgrund der neutralisierten Kriterien ist keine Unterteilung in invertierte und nicht-invertierte Items mehr nötig. Die relativen Häufigkeiten zeigen, dass die Skalenpunkte -3 und +3 von jeweils weniger als 10 % der Probanden ausgewählt wurden. Die übrigen fünf Skalenpunkte wurden vergleichbar häufig ausgewählt, wobei die Anteile zwischen 14 % und 18 % lagen. Anschließend wurde analysiert, bei wie vielen Items welche Skalenpunkte nie ausgewählt wurden. Dies war jedoch nie der Fall, d. h. von den Probanden wurden alle Skalenpunkte aller Kriterien genutzt. Die Verteilung war also im Vergleich zu den vorherigen Studien homogener.

Tabelle 37. Antwortverhalten der Probanden.

Skalenpunkt	Anteil aller Antworten [%]	Anteil Items an allen Items, bei denen Skalenpunkt von keinem Probanden ausgewählt wurde [%]
-3	9,56	0
-2	16,14	0
-1	18,48	0
0	17,25	0
+1	18,19	0
+2	14,40	0
+3	5,98	0

Objektivität, Reliabilität und Validität

Die Qualität eines Fragebogens lässt sich mithilfe von Kriterien der Testgüte bestimmen (Bortz & Döring, 2006). Die drei Hauptgütekriterien sind die Objektivität, Reliabilität und Validität (Lienert & Raatz, 1998). Auf sie soll nun genauer eingegangen werden.

Die Objektivität beschreibt die Unabhängigkeit der Ergebnisse einer Messung von der Person, die sie durchführt (Renner et al., 2012). Hierbei sind die Schritte der Durchführung, Auswertung und Interpretation zu beachten.

- Die Durchführungsobjektivität wird durch eine maximale Standardisierung der Durchführung erreicht (Bortz & Schuster, 2010). Essentiell hierfür ist es, Testmaterial zur Verfügung zu stellen. Darunter fallen neben dem Fragebogen auch schriftlich festgehaltene Instruktionen und eine Versuchsleiterinstruktion. Nach Schmidt-Atzert und Amelang (2012) ist die Objektivität dann hoch, wenn alle Einflussfaktoren festgelegt sind, die sich auf die Messung auswirken könnten. Bei der Bewertung des Lenkgefühls durch Normalfahrer ist die Fahrsituation, in der ein Bewertungskriterium evaluiert wird, ein bedeutender Einflussfaktor. Deshalb wurde im Versuchskonzept festgehalten, dass diese im Vorfeld festgelegt werden müssen. Ferner war die genaue Abfolge der Instruktionen innerhalb des Versuchsablaufs festgelegt. Außerdem waren alle Versuchsleiter instruiert, wie sie auf verschiedene Anfragen des Probanden reagieren sollten, beispielsweise bzgl. einer Pause oder der Wiederholung von Varianten.
- Die Auswertungsobjektivität beschreibt, dass die Vergabe von Testpunkten unabhängig vom Auswerter ist (Bortz & Döring, 2006). Zur Beantwortung des Fragebogens ist eine geschlossene, mit Punkten versehene Antwortskala vorgesehen. Da diese die reine Übertragung der Bewertungen der Probanden verlangt, ist nicht von Einschränkungen der Objektivität auszugehen.
- Die Interpretationsobjektivität ist dann gegeben, wenn alle Auswerter ein Testergebnis gleich interpretieren (Bühner, 2011). Da im Rahmen dieser Studie keine individuelle Einordnung bzw. Diagnostik der einzelnen Fahrer vorgenommen wurde, sind Einschränkungen der Objektivität auszuschließen.

Anschließend wurde die Reliabilität betrachtet. Diese beschreibt die Genauigkeit der Messung (Renner et al., 2012) und kann mit verschiedenen Methoden geschätzt werden. Da die Messgenauigkeit wesentlich durch die

Messbedingungen bestimmt wird, ist die Objektivität eine entscheidende Voraussetzung für die Reliabilität (Moosbrugger & Kelava, 2012).

- Die Paralleltest-Reliabilität und Testhalbierungsreliabilität basieren auf der Aufteilung des Fragebogens in zwei Hälften und dem Vergleich der erzielten Ergebnisse (Bortz & Döring, 2006; Bühner, 2011). Dieses Vorgehen war für diesen Anwendungsfall jedoch nicht geeignet. Schließlich wird nicht die Messung eines psychologischen Konstrukts angestrebt, sondern die Bewertung eines Sachverhalts. Deshalb war es nicht möglich, zwei Testhälften zu generieren, die inhaltlich die gleichen Aspekte abdecken und somit zu den gleichen Testwerten führen sollten.
- Die Retest-Reliabilität basiert auf der Wiedervorlage eines Tests und der berechneten Korrelation zwischen der ersten und erneuten Beantwortung (Schmidt-Atzert & Amelang, 2012). Zur Evaluation psychologischer Fragebögen wird häufig ein Intervall von mehreren Wochen zwischen beiden Erhebungszeitpunkten gewählt (Bortz & Döring, 2006). Dieses war im Rahmen der vorliegenden Arbeit in dieser Form weder realisierbar, noch sinnvoll. Vielmehr wurde darauf zurückgegriffen, dass pro Bewertungskriterium eine Variante wiederholt bewertet wurde (siehe auch Hypothese H3a). Somit wurden die ursprünglichen und wiederholten Bewertungen derselben Variante genutzt und Korrelationen zwischen ihnen berechnet. Die mittlere Korrelation über alle Items betrug für die Niveau-Bewertungen $r = .277$ und für die Gefallen-Bewertungen $r = .395$. Die mittlere Korrelation über alle Bewertungen lag demnach bei $r = .336$. Es zeigte sich dabei kein einheitliches Bild, ob bei bestimmten Kriterien allgemein höhere oder niedrigere Korrelationen vorlagen. Es war ebenfalls kein Unterschied zwischen den Korrelationen der unmittelbaren und getrennten Wiederholungen zu erkennen.
- Zur Bestimmung der internen Konsistenz wird häufig Cronbachs α herangezogen (Cronbach, 1951). Dieser Wert entspricht der gemittelten Testhalbierungsreliabilität eines Tests für alle möglichen Testhalbierungen (Bortz & Döring, 2006). Für die beiden Faktoren lag Cronbachs α bei .683. und .562. Diese Werte lassen in Abhängigkeit der Itemanzahlen und mittleren Korrelationen den Schluss zu, dass eine ausreichende interne Konsistenz vorliegt (Cortina, 1993).

Abschließend soll die Validität beleuchtet werden. Sie beschreibt den Grad der Genauigkeit, mit dem ein Test das misst, was er messen soll (Lienert & Raatz, 1998). Die Validität ist nach Schmidt-Atzert und Amelang (2012) das wichtigste Gütekriterium.

- Die Inhaltsvalidität beschreibt, inwiefern die Items eines Tests das zu messende Merkmal erfassen (Moosbrugger & Kelava, 2012). Da in der Regel keine empirische Überprüfung möglich ist, wird die Inhaltsvalidität häufig durch das Offenlegen der Vorgehensweise zur Entwicklung des Fragebogens belegt. Neben der Vorstellung des Konzepts und dem Berichten der einzelnen Schritte kann so auch über die Hinzunahme von Experten bei der Itemauswahl informiert werden (Schmidt-Atzert & Amelang, 2012). Demnach wurde der gesamte Entwicklungsprozess, der dem Fragebogen zugrunde liegt, in Kapitel 3.1.1 dargelegt. Abschließend muss für eine hohe Inhaltsvalidität sichergestellt werden, dass alle Aspekte des zu bewertenden Gegenstands abgedeckt werden (Bortz & Döring, 2006). Dies wurde in die Anforderungen an den Fragebogen aufgenommen (Kapitel 3.1.) und entsprechend beachtet.
- Die Kriteriumsvalidität beschreibt den Zusammenhang zwischen dem Ergebnis einer Messung und einem Kriterium außerhalb der Testsituation (Renner et al., 2012). Die Überprüfung der Kriteriumsvalidität gestaltet sich oftmals schwierig, da keine geeigneten Außenkriterien identifiziert werden können (Bortz & Döring, 2006). Dies trifft auch im vorliegenden Forschungsbereich zu. Für einzelne Items können Außenkriterien gefunden werden, bspw. der Lenkradwinkelbedarf für die *Lenkpräzision*. Es gibt jedoch keine messbaren Variablen, welche zur Validierung des gesamten Fragebogens geeignet sind.

- Unter der Konstruktvalidität versteht sich, dass ein Test das Konstrukt erfasst, was er erfassen soll (Schmidt-Atzert & Amelang, 2012). Hierzu wird die Stärke des Zusammenhangs mit weiteren Indikatoren des zu erfassenden Konstrukts bestimmt (konvergente Validität). Diese wird dann mit der Stärke des Zusammenhangs mit Indikatoren anderer Konstrukte (diskriminante Validität) verglichen. In diesem Zusammenhang wird häufig die Multitrait-Multimethod-Methode (Campbell & Fiske, 1959) angesprochen. Auch hier ist es essentiell, dass verwandte Konstrukte vorliegen. Dies ist jedoch für die subjektive Bewertung des Lenkgefühls nicht der Fall. Ein Indiz für Konstruktvalidität ist, dass die gewonnenen Messwerte gemäß den aus Theorie und Empirie abgeleiteten Hypothesen ausfallen (Bortz & Döring, 2006).

Die vorgestellten Erkenntnisse zu den Hauptgütekriterien Objektivität, Reliabilität und Validität zeigen, dass das vorliegende Versuchskonzept die Anforderungen erfüllte. Ferner sprechen die Ergebnisse der Faktoren-, Item- und Reliabilitätsanalysen für eine Eignung des Fragebogens. Deshalb wurde Hypothese 1, dass der Fragebogen eine ausreichende psychometrische Qualität aufweist, beibehalten.

4.3.2. Vorhersagbarkeit der Gefallen-Bewertungen in Abhängigkeit der Niveau-Bewertung

H2: Die Gefallen-Bewertungen lassen sich in Abhängigkeit von den Niveau-Bewertungen vorhersagen.

Zur Überprüfung dieser Hypothese wurden Regressionen berechnet. Für alle acht Kriterien wurden Regressionsmodelle gefunden, die die Gefallen-Bewertung signifikant vorhersagten (Tabelle 38 und Abbildung 31). Sowohl quadratische als auch lineare Modelle führten zu den besten Vorhersagen. Die Varianzaufklärung lag zwischen 15 % beim *Anlenkmoment* und 84 % beim *Rückstellverhalten*. Die Regressionskoeffizienten sind in Anhang X zu finden.

Tabelle 38. Zusammenfassung der Regressionsmodelle zur Vorhersage der Niveau-Bewertung der acht Bewertungskriterien auf Basis der Gefallen-Bewertungen.

Bewertungskriterium	Art der Regression	<i>df</i>	<i>F</i>	<i>p</i>	korrr. <i>R</i> ²
Lenkmoment (Rangieren)	quadratisch	2, 441	310,15	< .001	.583
Lenkmoment (Slalom)	quadratisch	2, 441	48,66	< .001	.177
Anlenkmoment	quadratisch	2, 447	39,23	< .001	.145
Haltemoment	quadratisch	2, 447	133,93	< .001	.372
Rückstellverhalten	linear	1, 448	2398,16	< .001	.842
Lenkpräzision	linear	1, 447	2205,59	< .001	.831
Fahrbahnkontakt	quadratisch	2, 442	796,19	< .001	.782
Mittenzentrierung	linear	1, 442	510,48	< .001	.535

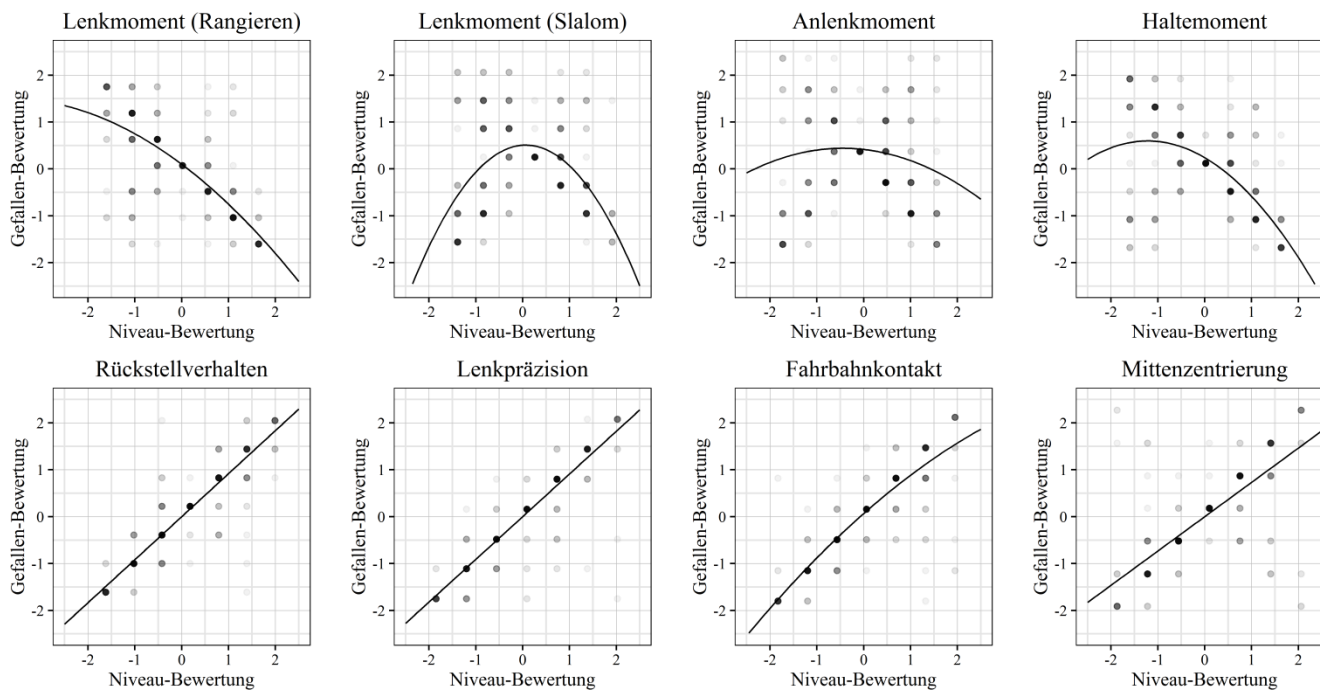


Abbildung 31. Punktdiagramme und berechnete Regressionskurven der Gefallen-Bewertungen der acht Bewertungskriterien in Abhängigkeit der Niveau-Bewertungen. Die Sättigung der Punkte nimmt mit größer werdender Fallzahl zu.

Hypothese 2, dass sich die Gefallen-Bewertung basierend auf den Niveau-Bewertungen vorhersagen lassen, wurde somit beibehalten.

4.3.3. Zuverlässigkeit der Bewertung

H3: Die wiederholte Bewertung der gleichen Variante der Lenkunterstützung führt nicht zu unterschiedlichen Bewertungen.

Die Hypothese wird wie folgt in zwei Unterhypothesen aufgeteilt.

H3a: Die wiederholte Bewertung der gleichen Variante führt nicht zu unterschiedlichen Niveau- und Gefallen-Bewertungen (Re-Test).

Wie im Versuchsdesign (Kapitel 4.2.1) gezeigt, wurden die Varianten ohne das Wissen der Probanden wiederholt. Die wiederholten Varianten wurden entweder an 3. und 4. Stelle (unmittelbar) oder an 1. und 6. Stelle (getrennt) innerhalb eines Kriteriums bewertet. Da anstatt eines vollständig randomisierten Versuchsplans vier Sets verwendet wurden, wurden nicht alle Varianten in jedem Manöver sowohl unmittelbar, als auch getrennt wiederholt.

Mithilfe von t -Tests für verbundene Stichproben wurden die regulären mit den wiederholten Bewertungen verglichen. Insgesamt wurden für die Niveau- und Gefallen-Bewertungen jeweils 23 Paarvergleiche durchgeführt. Von den Vergleichen waren gemäß der Definition elf unmittelbar und zwölf getrennt. Bis auf zwei Vergleiche gab es keine signifikanten Unterschiede zwischen den Paaren. Zum einen war die Gefallen-Bewertung von V5 beim *Haltemoment* bei der unmittelbaren Wiederholung signifikant höher, $t(19) = -2.90$, $p = .023$, $d = 1.08$. Weiterhin war die Gefallen-Bewertung der *Mittenzentrierung* der V4 bei der getrennten

Wiederholung signifikant höher, $t(36) = -3,17$, $p = .008$, $d = 1.01$. Die deskriptiven Statistiken sind in Anhang XI zu finden.

Aufgrund des Versuchsdesigns lag bei einigen Vergleichen lediglich eine Stichprobengröße von $N = 18$ vor. Da dies die Ursache für ein nicht-signifikantes Ergebnis sein könnte, wurden auch die vier nicht signifikanten Ergebnisse mit mittleren Effektstärken einer Analyse unterzogen. Bei der *Mittenzentrierung* lag für einen Vergleich eine mittlere Effektstärke vor, bei der *Lenkpräzision* für drei. In drei Fällen unterschied sich die Ausprägung, in einem das Gefallen. Es ergab sich jedoch kein einheitliches Muster, ob die erste oder wiederholte Bewertung tendenziell besser ausfiel. Auffällig war jedoch, dass in allen vier Fällen die Varianten getrennt wiederholt wurden, d. h. an erster und letzter Position innerhalb des Kriteriums.

Nachdem nur zwei signifikante Unterschiede vorlagen, wurde Hypothese 3a, dass die wiederholte Bewertung der gleichen Variante nicht zu unterschiedlichen Niveau- und Gefallen-Bewertungen führt, beibehalten.

H3b: Die Niveau- und Gefallen-Bewertungen zeigen keinen Unterschied zwischen der Referenz und der blind als Variante gefahrenen Referenz an.

Wie in Kapitel 4.2.3 beschrieben, wurde V3 gleichzeitig als Referenz genutzt. Da dies den Probanden jedoch nicht mitgeteilt wurde, konnte geprüft werden, ob ein Unterschied zwischen den Bewertungen der Referenz und der blind als Variante gefahrenen Referenz besteht. Hierzu wurden die nicht z -transformierten Niveau- und Gefallen-Bewertungen der acht Kriterien herangezogen. Die Bewertungen der V3 (Referenz) wurden mit t -Tests bei einer Stichprobe gegen den Wert 0 getestet, da dieser auf der Skala für keinen wahrgenommenen Unterschied stand. Für das *Rückstellverhalten* ergab sich bei der Niveau-Bewertung ein signifikant höherer Wert, $t(74) = 2,45$, $p = .041$, $d = .28$. Auch die Gefallen-Bewertung war vor der Korrektur signifikant größer als 0, $t(74) = 2,21$, $p = .071$, $d = .26$. Die deskriptiven Statistiken sind in Anhang XII zu finden.

Somit wurde Hypothese 3b, dass die Niveau- und Gefallen-Bewertungen der Referenz und der blind als Variante gefahrenen Referenz nicht verschieden sind, mit Ausnahme des Kriteriums Rückstellverhalten beibehalten.

Nachdem beide Unter-Hypothesen beibehalten wurden, wurde auch Hypothese 3, dass die wiederholte Bewertung der gleichen Variante der Lenkunterstützung nicht zu unterschiedlichen Bewertungen führt, beibehalten.

4.3.4. Wahrnehmbarkeit der Unterschiede zwischen den Varianten

H4: Die objektiv vorhandenen Unterschiede in der Höhe des aufzuwendenden Lenkmoments spiegeln sich in entsprechend unterschiedlichen Niveau-Bewertungen wider.

Zur Überprüfung dieser Hypothese wurden ANOVAs mit Messwiederholung (GLM4) und anschließende wiederholte Kontraste berechnet.

Ferner wurde analysiert, wie häufig die Probanden den zugrunde liegenden Unterschied korrekt, nicht oder inkorrekt wahrnahmen. Hierzu wurden auf Basis der Rohdaten die prozentualen Häufigkeiten ausgezählt (Tabelle 39). Für den Unterschied zwischen V3 (3,4 Nm) und V4 (4,1 Nm) mit beispielhaften Niveau-Bewertungen zu *Lenkmoment (Rangieren)* bedeutet dies:

- Der objektiv vorhandene Unterschied wurde korrekt wahrgenommen: Niveau₃ = +1, Niveau₄ = +2

- Es wurde kein Unterschied wahrgenommen: Niveau₃ = +2, Niveau₄ = +2
- Es wurde ein entgegengesetzter, inkorrekt wahrgenommen: Niveau₃ = +3, Niveau₄ = +2

Tabelle 39. Prozentuale Häufigkeiten der korrekt wahrgenommenen, nicht wahrgenommenen und inkorrekt wahrgenommenen Unterschiede zwischen den fünf Varianten für die Kriterien *Lenkmoment (Rangieren)*, *Lenkmoment (Slalom)*, *Anlenkmoment* und *Haltemoment*.

Verglichene Varianten [Differenz in Nm]	Wahrgenommener Unterschied	Lenkmoment (Rangieren)	Lenkmoment (Slalom)	Anlenkmoment	Haltemoment	Mittelwert
V1 vs. V2 [1 Nm]	korrekt	78,7	80,0	82,7	76,0	79,3
	kein	16,0	18,7	10,7	18,7	16,0
	inkorrekt	5,3	1,3	6,7	5,3	4,7
V2 vs. V3 [0,7 Nm]	korrekt	68,0	66,7	77,3	88,0	75,0
	kein	18,7	24,0	14,7	8,0	16,3
	inkorrekt	13,3	9,3	8,0	4,0	8,7
V3 vs. V4 [0,7 Nm]	korrekt	86,7	82,7	88,0	73,3	82,7
	kein	10,7	10,7	9,3	22,7	13,3
	inkorrekt	2,7	6,7	2,7	4,0	4,0
V4 vs. V5 [0,4 Nm]	korrekt	74,7	57,3	70,7	56,0	64,7
	kein	22,7	36,0	24,0	14,7	24,3
	inkorrekt	2,7	6,7	5,3	29,3	11,0

Der Vergleich der Bewertungen ergab, dass durchschnittlich 79 % der Fahrer die Zunahme im maximalen Lenkmoment von 1 Nm von V1 zu V2 bemerkten. V3 wurde bei einem Unterschied von 0,7 Nm von 75 % der Fahrer als schwergängiger als V2 bewertet. Ebenfalls 0,7 Nm Unterschied lagen zwischen V3 und V4, was bei 83 % der Fahrer zu höheren Bewertungen führte. Schlussendlich verlangte V5 0,4 Nm mehr Lenkmoment als V4, was im Mittel von 65 % der Fahrer entsprechend wahrgenommen wurde. Mit 88 % korrekt entdecktem Unterschied war beim *Anlenkmoment* und *Haltemoment* die größte Aufdeckungsrate zu verzeichnen. Mit 56 % korrekt entdecktem Unterschied lag beim *Haltemoment* gleichzeitig der geringste Anteil vor. Außerdem war mit 29,3 % die höchste Rate der falsch detektierten Unterschiede festzustellen.

Es ist hervorzuheben, dass V2 und V4 innerhalb des Manövers jeweils direkt im Anschluss an die Referenz, welche V3 entsprach, bewertet werden konnten. Somit war hier ein unmittelbarer Vergleich möglich. Die Unterschiede zwischen V1 und V2 sowie V4 und V5 wurden indirekt aus den vergebenen Ratings in Bezug auf die Referenz berechnet.

Nachfolgend werden die Ergebnisse der Varianzanalysen mit Messwiederholung vorgestellt. Die deskriptiven Statistiken sind in Anhang XIII zu finden. Für alle vier Kriterien *Lenkmoment (Rangieren)*, *Lenkmoment (Slalom)*, *Anlenkmoment* und *Haltemoment* lag ein signifikanter Haupteffekt vor, d. h. die Niveau-Bewertungen waren signifikant verschieden (Tabelle 40 und Abbildung 32). Die angeschlossenen wiederholten Kontraste waren ebenfalls in allen Fällen signifikant. Es lagen bis auf eine Ausnahme große Effektstärken zugrunde.

Tabelle 40. Ergebnisse der Varianzanalysen mit Messwiederholung zum Vergleich der Niveau-Bewertungen der fünf Varianten in den Kriterien *Lenkmoment (Rangieren)*, *Lenkmoment (Slalom)*, *Anlenkmoment* und *Haltemoment*.

Bewertungskriterium	df _{Effekt}	df _{Fehler}	<i>F</i>	<i>p</i>	partielles η^2	<i>r</i>
Lenkmoment (Rangieren)	3,39	247,15	436,43	< .001	.86	
V1 vs. V2	1	73	89,47	< .001		.74
V2 vs. V3	1	73	48,18	< .001		.63
V3 vs. V4	1	73	127,37	< .001		.80
V4 vs. V5	1	73	91,61	< .001		.75
Lenkmoment (Slalom)	3,56	259,85	411,80	< .001	.85	
V1 vs. V2	1	73	130,89	< .001		.80
V2 vs. V3	1	73	62,87	< .001		.68
V3 vs. V4	1	73	99,87	< .001		.76
V4 vs. V5	1	73	44,35	< .001		.61
Anlenkmoment	3,76	278,57	459,09	< .001	.86	
V1 vs. V2	1	74	68,900	< .001		.69
V2 vs. V3	1	74	95,157	< .001		.75
V3 vs. V4	1	74	154,968	< .001		.82
V4 vs. V5	1	74	81,134	< .001		.72
Haltemoment	4	296	365,05	< .001	.83	
V1 vs. V2	1	74	88,237	< .001		.74
V2 vs. V3	1	74	164,160	< .001		.83
V3 vs. V4	1	74	82,254	< .001		.73
V4 vs. V5	1	74	12,641	.002		.38

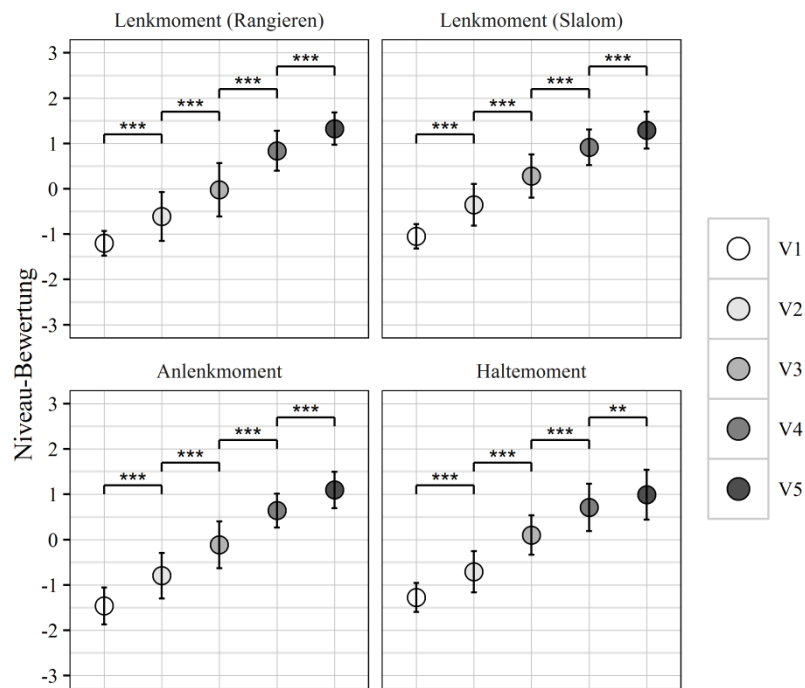


Abbildung 32. Mittelwerte der Niveau-Bewertungen der fünf Varianten in den Kriterien *Lenkmoment (Rangieren)*, *Lenkmoment (Slalom)*, *Anlenkmoment* und *Haltemoment*. Die Fehlerbalken repräsentieren die Standardabweichung.

Somit wurde Hypothese 4, dass die objektiv vorhandenen Unterschiede in der Höhe des aufzuwendenden Lenkmoments detektiert werden, beibehalten.

4.3.5. Vorhersagbarkeit der Niveau-Bewertung in Abhängigkeit der Variante

H5: Die Niveau-Bewertungen der direkt vom aufzuwendenden Lenkmoment abhängigen Bewertungskriterien lassen sich in Abhängigkeit von Parametern der Lenkungsauslegung vorhersagen.

Zur Überprüfung dieser Hypothese wurden Regressionen berechnet. Das maximal vom Fahrer aufzuwendende Lenkmomentniveau wurde als Prädiktor herangezogen. Für alle fünf Kriterien wurden Regressionsmodelle gefunden, die die Gefallen-Bewertung signifikant vorhersagten (Tabelle 41 und Abbildung 33). Dabei lag die Varianzaufklärung zwischen 77 % bei *Lenkmoment (Rangieren)* und 83 % bei *Lenkmoment (Slalom)*. Deutlich geringer als bei allen anderen Kriterien war der Zusammenhang mit der Niveau-Bewertung des *Rückstellverhaltens*. Dieses konnte nur zu 9 % vorhergesagt werden. In allen Fällen führte ein quadratisches Modell zu den besten Vorhersagen. Die Regressionskoeffizienten sind in Anhang XIV zu finden.

Hypothese 5 wurde beibehalten, da die Niveau-Bewertung der direkt vom aufzuwendenden Lenkmoment abhängigen Bewertungskriterien basierend auf den Parametern der Lenkungsauslegung vorhergesagt werden können.

Tabelle 41. Zusammenfassung der Regressionsmodelle zur Vorhersage der Niveau-Bewertung in den Kriterien *Lenkmoment (Rangieren)*, *Lenkmoment (Slalom)*, *Anlenkmoment*, *Haltemoment* und *Rückstellverhalten* auf Basis des maximalen Lenkmoments der Varianten.

Bewertungskriterium	Art der Regression	<i>df</i>	<i>F</i>	<i>p</i>	korrr. <i>R</i> ²
Lenkmoment (Rangieren)	quadratisch	2, 441	740,10	< .001	.769
Lenkmoment (Slalom)	quadratisch	2, 441	1095,58	< .001	.832
Anlenkmoment	quadratisch	2, 447	898,28	< .001	.800
Haltemoment	quadratisch	2, 447	839,16	< .001	.789
Rückstellverhalten	quadratisch	2, 447	23,92	< .001	.093

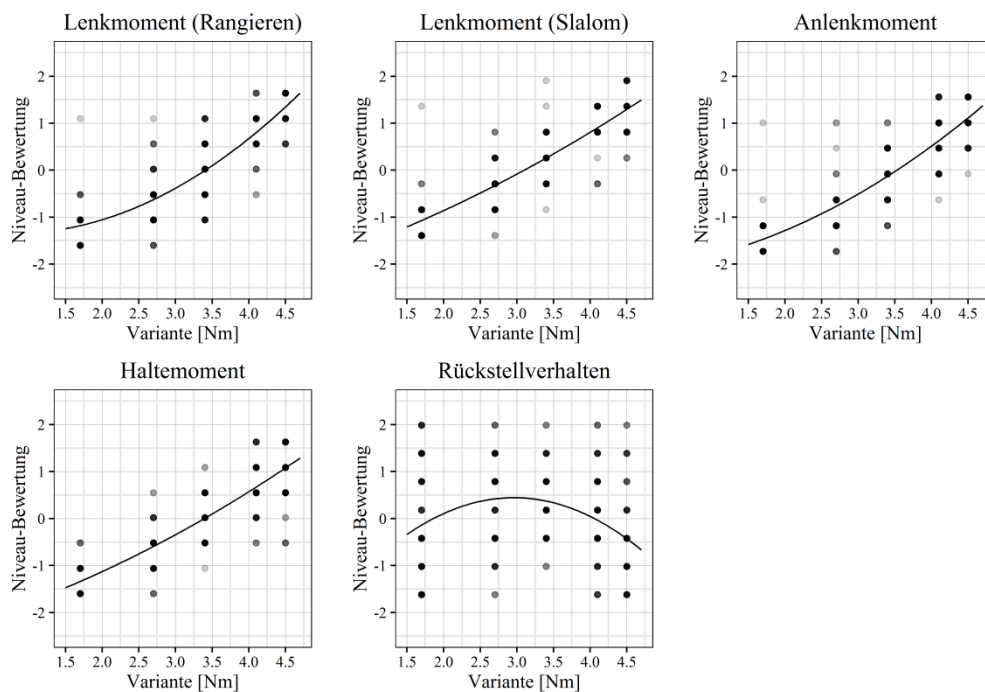


Abbildung 33. Punktdiagramme und berechnete Regressionskurven der Niveau-Bewertungen in den Kriterien *Lenkmoment (Rangieren)*, *Lenkmoment (Slalom)*, *Anlenkmoment*, *Haltemoment* und *Rückstellverhalten* in Abhängigkeit des maximalen Lenkmoments der Varianten. Die Sättigung der Punkte nimmt mit größer werdender Fallzahl zu.

4.3.6. Vorhersagbarkeit der Gefallen-Bewertungen in Abhängigkeit der Variante

H6: Die Gefallen-Bewertungen der direkt vom aufzuwendenden Lenkmoment abhängigen Bewertungskriterien lassen sich in Abhängigkeit von Parametern der Lenkungsauslegung vorhersagen.

Zur Überprüfung dieser Hypothese wurden Regressionen berechnet. Das maximal vom Fahrer aufzuwendende Lenkmomentniveau wurde als Prädiktor herangezogen. Für alle fünf Kriterien wurden Regressionsmodelle gefunden, die die Gefallen-Bewertung signifikant vorhersagten (Tabelle 42 und Abbildung 34). Dabei lag die Varianzaufklärung zwischen 14 % beim *Lenkmoment (Slalom)* und 45 % beim *Lenkmoment (Rangieren)*. Der Zusammenhang mit der Gefallen-Bewertung des *Rückstellverhaltens* war deutlich niedriger, was sich in der

erklärten Varianz von 9 % widerspiegelt. In allen Fällen führten quadratische Modelle zu den besten Vorhersagen. Die Regressionskoeffizienten sind in Anhang XV zu finden.

Tabelle 42. Zusammenfassung der Regressionsmodelle zur Vorhersage der Gefallen-Bewertung in den Kriterien *Lenkmoment (Rangieren)*, *Lenkmoment (Slalom)*, *Anlenkmoment*, *Haltemoment* und *Rückstellverhalten* auf Basis des maximalen Lenkmoments der Varianten.

Kriterium	Art der Regression	<i>df</i>	<i>F</i>	<i>p</i>	kor. <i>R</i> ²
Lenkmoment (Rangieren)	quadratisch	2, 441	183,54	< .001	.452
Lenkmoment (Slalom)	quadratisch	2, 441	35,64	< .001	.135
Anlenkmoment	quadratisch	2, 447	50,62	< .001	.181
Haltemoment	quadratisch	2, 447	94,45	< .001	.294
Rückstellverhalten	quadratisch	2, 447	22,29	< .001	.087

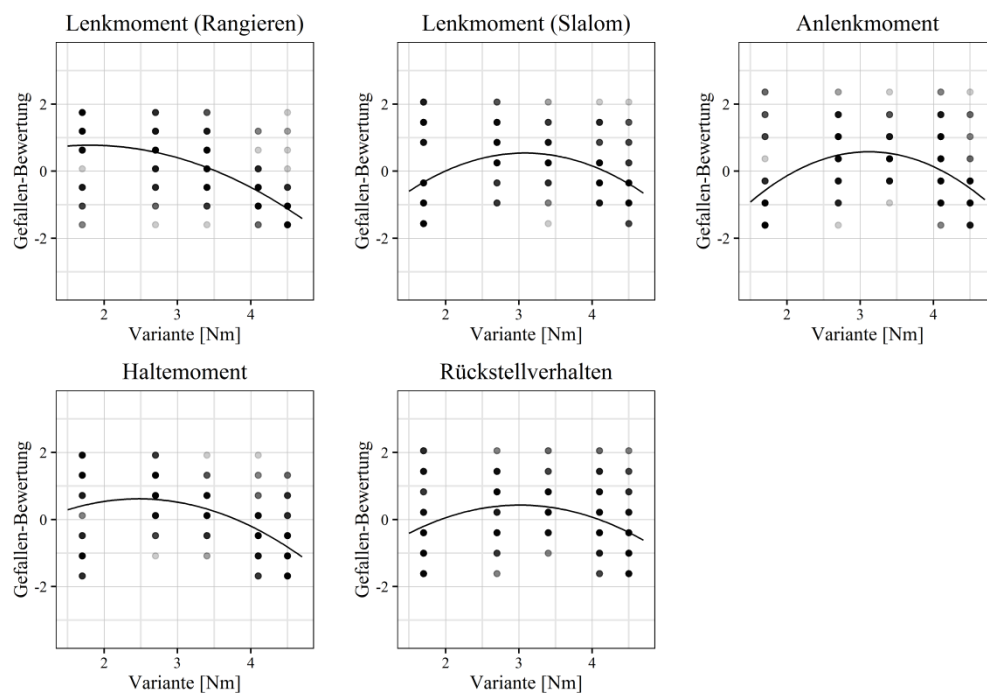


Abbildung 34. Punktdiagramme und berechnete Regressionskurven der Gefallen-Bewertungen in den Kriterien *Lenkmoment (Rangieren)*, *Lenkmoment (Slalom)*, *Anlenkmoment*, *Haltemoment* und *Rückstellverhalten* in Abhängigkeit des maximalen Lenkmoments der Varianten. Die Sättigung der Punkte nimmt mit größer werdender Fallzahl zu.

Hypothese 6 wurde beibehalten, da die Gefallen-Bewertung der direkt vom aufzuwendenden Lenkmoment abhängigen Bewertungskriterien basierend auf den Parametern der Lenkungsauslegung vorhergesagt werden können.

4.3.7. Explorative Analysen

EXP: Unterscheidet sich die Zuverlässigkeit der Bewertungen in Abhängigkeit der Gesamtfahrerfahrung und des Alters?

Zur weiteren Verfeinerung des Versuchskonzepts sollte der Einfluss der demographischen Daten auf die Zuverlässigkeit der Bewertungen untersucht werden. Deshalb wurden die Berechnungen zu H3b (Kapitel 4.3.3) in Abhängigkeit der Fahrerfahrung der Probanden wiederholt. Um eine entsprechende Stichprobengröße zu erhalten, wurden alle Werte gemeinsam analysiert. Es wurden also die nicht z-transformierten Niveau- und Gefallen-Bewertungen der acht Kriterien gegen den Wert 0 getestet, da dieser auf der Skala für keinen wahrgenommenen Unterschied stand. Die deskriptiven Statistiken sind in Anhang XVI zu finden.

Im ersten Schritt wurden die Analysen der Fahrerfahrung durchgeführt. Die Fahrer wurden in folgende drei Gruppen eingeteilt:

- Geringe Fahrerfahrung: Gesamt-Fahrleistung bis 600.000 km ($n = 24$)
- Mittlere Fahrerfahrung: Gesamt-Fahrleistung zwischen 600.001 und 1.999.999 km ($n = 22$)
- Hohe Fahrerfahrung: Gesamt-Fahrleistung ab 2.000.000 km ($n = 27$)

Die t -Tests bei einer Stichprobe zeigten, dass die Bewertungen der Fahrer mit niedriger Fahrerfahrung nicht signifikant verschieden von 0 waren, $t(379) = 0,54$, $p = .590$, $d = .03$. Die Bewertungen der Fahrer mit mittlerer ($t(351) = 2,97$, $p = .009$, $d = .16$) und hoher Fahrerfahrung ($t(429) = 2,65$, $p = .017$, $d = .13$) waren jedoch signifikant größer als 0. In beiden Fällen lag die Effektstärke unterhalb der Grenze für einen kleinen Effekt. Somit schätzten die Fahrer mit mittlerer und hoher Fahrerfahrung die als Variante getarnte Referenz bezüglich des Niveaus und Gefallens signifikant höher ein als die Referenz selbst.

Anschließend wurden die Analysen des Alters durchgeführt. Die Fahrer wurden in folgende drei Gruppen eingeteilt:

- Jüngere Fahrer: Alter bis 40 Jahre ($n = 26$)
- Fahrer mittleren Alters: Alter zwischen 41 und 50 Jahre ($n = 25$)
- Ältere Fahrer: Alter über 50 Jahre ($n = 24$)

Die t -Tests bei einer Stichprobe zeigten, dass die Bewertungen der jungen Fahrer signifikant größer als 0 waren, $t(411) = 3,17$, $p = .010$, $d = .16$. Die Effektstärke lag unterhalb der Grenze für einen kleinen Effekt. Für die Fahrer mittleren Alters, $t(397) = 1,35$, $p = .266$, $d = .07$, und die älteren Fahrer, $t(381) = 1,22$, $p = .267$, $d = .06$, lagen keine Unterschiede vor. Somit schätzten die jüngeren Fahrer die als Variante getarnte Referenz bezüglich des Niveaus und Gefallens signifikant höher ein als die Referenz selbst.

Insgesamt zeigten die Analysen somit Einschränkungen der Zuverlässigkeit der Bewertungen bei jungen Fahrern sowie Fahrern mit mittlerer bzw. hoher Fahrerfahrung auf. In allen Fällen lagen jedoch Effektstärken unter der Grenze für einen kleinen Effekt vor.

4.3.8. Ableitung der Gestaltungsempfehlungen

Es wurden Varianzanalysen mit Messwiederholung für die Gefallen-Bewertungen aller acht Kriterien berechnet (GLM 4). Die signifikanten Haupteffekte aller Kriterien zeigten, dass sich die Gefallen-Bewertungen der Varianten insgesamt unterscheiden (Tabelle 43). Die deskriptiven Statistiken sind in Anhang XVII zu finden.

Tabelle 43. Ergebnisse der Varianzanalysen mit Messwiederholung zum Vergleich der Gefallen-Bewertungen der fünf Varianten in allen Bewertungskriterien.

Kriterium	df_{Effekt}	df_{Fehler}	F	p	partiell η^2
Lenkmoment (Rangieren)	2,53	184,41	81,42	< .001	.53
Lenkmoment (Slalom)	3,24	236,57	16,42	< .001	.18
Anlenkmoment	2,74	202,83	26,12	< .001	.26
Haltemoment	2,31	170,93	39,56	< .001	.35
Rückstellverhalten	3,32	245,46	12,52	< .001	.15
Lenkpräzision	2,83	206,66	19,29	< .001	.21
Fahrbahnkontakt	3,25	233,74	36,49	< .001	.34
Mittenzentrierung	3,27	238,41	12,64	< .001	.15

Die Ergebnisse der im Anschluss berechneten paarweisen Vergleiche werden nun vorgestellt (Tabelle 44). Aus ihnen werden die für das jeweilige Kriterium bevorzugten Varianten abgeleitet. Eine Visualisierung ist in Abbildung 35 zu finden. Auf eine Darstellung der signifikanten Effekte wird aufgrund ihrer großen Zahl verzichtet. Vielmehr werden die Kennwerte der Einzelvergleiche ebenfalls in Anhang XVII berichtet.

In der Rangiersituation wurden bezüglich des *Lenkmoments* V1 und V2 mit den geringsten aufzuwendenden Lenkmomenten bevorzugt. Im Oval wurden zwei Bewertungskriterien erfasst. Hinsichtlich des *Haltemoments* erzielten V1 und V2 mit den geringsten aufzuwendenden Lenkmomenten die höchsten Werte beim Gefallen. Bei V1 fällt die vergleichsweise große Standardabweichung auf, welche für sehr unterschiedliche Gefallen-Bewertungen der Fahrer spricht. Hinsichtlich des *Rückstellverhaltens* gefielen den Fahrern die V2 und V3 am besten, welche sich durch die höchsten Rückstellmomente auszeichneten. Beim Slalom wurden ebenfalls zwei Kriterien bewertet. Hier erzielten die V2 und V3 bezüglich des *Lenkmoments* die höchsten Gefallen-Bewertungen. Sie zeichneten sich durch niedrige bis mittlere aufzuwendende Lenkmomente aus. Bezüglich des *Fahrbahnkontakts* schnitten V3, V4 und V5 am besten ab. Somit führten Varianten mit mittlerem bis hohem aufzuwendendem Lenkmoment zu vergleichbaren Gefallen-Bewertungen. Beim Spurwechsel waren die Gefallen-Bewertungen des *Anlenkmoments* für V2 und V3 mit mittlerem aufzuwendenden Lenkmoment am höchsten. Die *Lenkpräzision* wurde für alle Varianten außer der leichtgängigsten V1 vergleichbar eingeschätzt. In dieser Fahrsituation erhielten also sowohl Varianten mit mittlerem als auch mit hohem Lenkmomentniveau hohe Gefallen-Bewertungen. Bei der Geradeausfahrt wurde die *Mittenzentrierung* bewertet. Bei dieser gefielen den Probanden V3, V4 und V5 am besten.

Tabelle 44. Gestaltungsempfehlungen in Abhängigkeit der Fahrsituation.

Fahrsituation und korrespondierende Bewertungskriterien	Bevorzugte Varianten	Relevante Kennwerte
Rangieren		
Lenkmoment	V1, V2	Lenkmoment: 1,7 Nm, 2,7 Nm
Oval		
Haltemoment	V1, V2	
Rückstellverhalten	V2, V3	Lenkmoment: 1,7 Nm, 2,7 Nm Lenkmoment: 2,7 Nm, 3,4 Nm / Rückstellmoment: 4,3 Nm, 3,5 Nm
Slalom		
Lenkmoment	V2, V3	Lenkmoment: 2,7 Nm, 3,4 Nm
Fahrbahnkontakt	V3, V4, V5	Lenkmoment: 3,4 Nm, 4,1 Nm, 4,5 Nm
Spurwechsel		
Anlenkmoment	V2, V3	Lenkmoment: 2,7 Nm, 3,4 Nm
Lenkpräzision	V2, V3, V4, V5	Lenkmoment: 2,7 Nm, 3,4 Nm, 4,1 Nm, 4,5 Nm
Geradeausfahrt		
Mittenzentrierung	V3, V4, V5	Lenkmoment: 3,4 Nm, 4,1 Nm, 4,5 Nm

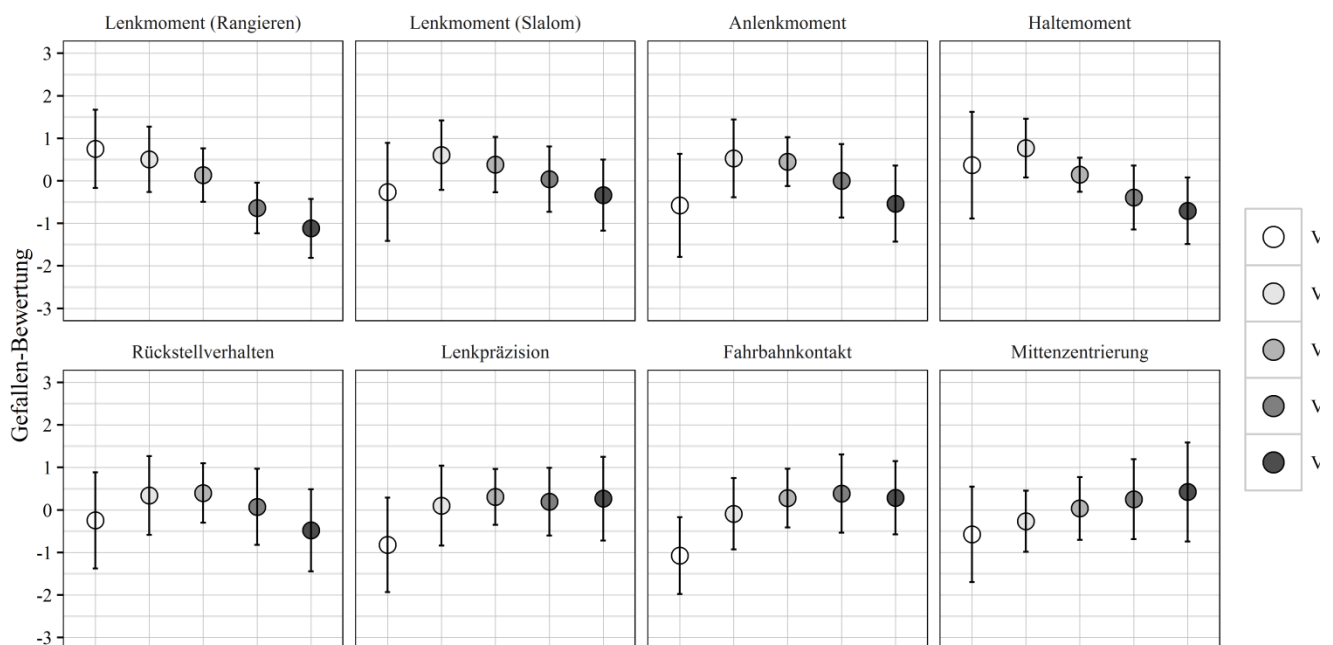


Abbildung 35. Mittelwerte der Gefallen-Bewertungen der fünf Varianten in allen Bewertungskriterien. Die Fehlerbalken repräsentieren die Standardabweichung. Auf eine Darstellung der signifikanten Effekte wird aus Gründen der Übersichtlichkeit verzichtet.

4.4. Diskussion

4.4.1. Summative Evaluation des Versuchskonzepts

H1: Der Fragebogen weist eine ausreichende psychometrische Qualität auf.

Zur Überprüfung dieser Hypothese wurde eine Faktoren-, Item- und Reliabilitätsanalyse durchgeführt. Ferner wurde der Fragebogen in Hinblick auf die Gütekriterien beleuchtet.

Faktoren-, Item- und Reliabilitätsanalyse

Die Hauptkomponentenanalyse zeigte eine zweifaktorielle Lösung an, welche insgesamt 56 % der Varianz erklärte. Der Faktor *Fahrzeugreaktion* beinhaltete die fünf Kriterien *Lenkmoment (Rangieren)*, *Lenkmoment (Slalom)*, *Anlenkmoment*, *Haltemoment* und *Rückstellverhalten*. Er erklärte 32 % der Varianz und wies ein Cronbachs α von .705 auf. Der Faktor *Lenkaktivität* beinhaltete die Kriterien *Fahrbahnkontakt*, *Lenkpräzision* und *Mittenzentrierung* und erklärte die übrigen 24 % der Varianz. Cronbachs α betrug .596. Die internen Konsistenzen können bei dieser kleinen Anzahl an Items als zufriedenstellend eingeschätzt werden.

Im Vergleich zu den vorherigen Studien deckte die Verteilung der Antworten über alle Skalenpunkte hinweg keine Antwortstile auf. Es lag im Gegensatz zu den vorherigen Studien keine Häufung an den Polen vor. Außerdem wurde bei allen Items jeder Skalenpunkt von mehreren Probanden ausgewählt. Dies spricht zum einen dafür, dass mit den Varianten das gesamte Spektrum guter und schlechter Auslegungen abgebildet wurde. Andererseits zeigt das Fehlen von Skalenpunkten, die von keinem Probanden gewählt wurden, dass die Anzahl an Skalenpunkten nicht zu groß war. Insgesamt spiegelt sich hierin maßgeblich der Erfolg der zwischen Studie II und III vorgenommenen Veränderungen des Versuchskonzepts wider.

Weiterhin soll auf die Varianzaufklärung von 56 % eingegangen werden. Zur Einordnung der gefundenen Höhe der Varianzaufklärung liegen keine allgemeinen Richtwerte vor. Da auch die übrigen Studien zum Lenkgefühl keine Angaben hierzu machen, wird zum Vergleich auf themenverwandte Arbeiten zurückgegriffen. Jürgensohn und Kolrep (2006) entwickelten einen Fragebogen zur subjektiven Bewertung des Spiegelzitterns. Mit den verwendeten zehn semantischen Differenzialen erzielten sie eine Varianzaufklärung von 60 %. Für einen Fragebogen zur Erfassung von Fahrkomfort und Fahrspaß mit 41 Items wurde eine Varianzaufklärung von 64 % berichtet (Engelbrecht, 2013). Eine vergleichbare Varianzaufklärung (63 %) wurde bei der subjektiven Bewertung auditiver Wahrnehmungen im Fahrzeuginnenraum unter Verwendung von 49 semantischen Differenzialen gefunden (Hempel, 2001). Der Anteil der aufgeklärten Varianz liegt somit leicht über der, die für den Fragebogen zum Lenkgefühl gefunden wurde. Im Vergleich zu diesen Studien besteht der vorliegende Fragebogen aber aus einer zum Teil erheblich kleineren Itemanzahl. Ferner ist beispielsweise im Vergleich zur ersten Studie bzgl. der Bewertung des Spiegelzitterns von einer stärkeren Individualität der Präferenzen auszugehen. Somit ist die Varianzaufklärung insgesamt als zufriedenstellend einzustufen.

Ferner wird die mittlere Inter-Korrelation zwischen den Gefallen-Bewertungen eines Faktors diskutiert. Diese lag bei beiden Faktoren knapp über .3 und war somit niedriger als bei den Daten aus Studie II. Dies liegt vermutlich darin begründet, dass die Bewertungen der Fahrer in diesem Versuch über die gesamte Skala verteilt waren und

somit insgesamt mehr Varianz in den Daten vorlag. Die gefundene Stärke der Korrelation entspricht außerdem dem von Rothhämel et al. (2011a) berichteten Zusammenhang zwischen den Niveau-Bewertungen von fünf Kriterien. Im Gegensatz hierzu wird für einen Fragebogen zur Lenkungsbewertung mit sieben Items eine mittlere Inter-Item-Korrelation von .7 berichtet (Riedel & Arbinger, 1997). Bei Decker (2009) lag die mittlere Korrelation zwischen den Gefallen-Bewertungen ebenfalls in diesem Bereich. Andere Studien finden sogar Korrelationen von bis zu .9 (Anand, 2014). Höhere Korrelationen zwischen den Items sprechen jedoch nicht zwangsläufig für eine höhere Qualität des Fragebogens. Aus derart starken Zusammenhängen kann geschlossen werden, dass die angebotenen Varianten über alle Kriterien hinweg entweder gut oder schlecht bewertet wurden. Sie können auch dadurch zustande kommen, dass die Kriterien untereinander nicht differenzierbar sind. Ferner können mehreren Items eng verwandte bzw. identische Konstrukte zugrunde liegen, wie es beispielsweise in Studie II bei den Sub-Kriterien zum *Rückstellverhalten* und zum *Korrekturbedarf* der Fall war.

Insgesamt sind die Ergebnisse aus der explorativen Faktorenanalyse sowie der Item- und Reliabilitätsanalyse als zufriedenstellend einzustufen.

Objektivität, Reliabilität und Validität

Im Folgenden soll auf die Gütekriterien Objektivität, Reliabilität und Validität eingegangen werden.

Hinsichtlich der Objektivität werden die Schritte Durchführung, Auswertung und Interpretation beachtet (Renner et al., 2012). Die Durchführungsobjektivität wurde durch eine maximale Standardisierung erreicht, beispielsweise durch schriftlich festgehaltene Instruktionen. Die Auswertungsobjektivität ist aufgrund der verwendeten, geschlossenen Skala gegeben. Die Interpretationsobjektivität ist ebenfalls nicht eingeschränkt, da keine Einzelfallbetrachtung individueller Bewertungen vorgenommen wurde. Somit war hinsichtlich der Objektivität nicht von Limitationen auszugehen.

Die Reliabilität beschreibt die Genauigkeit der Messung (Renner et al., 2012). Im Rahmen dieses Versuchs wurde die Retest-Reliabilität überprüft (Schmidt-Atzert & Amelang, 2012). Hierzu wurden die ursprünglichen und wiederholten Bewertungen derselben Variante genutzt. Die durchschnittliche Korrelation lag für die Niveau-Bewertungen bei $r = .277$, für die Gefallen-Bewertungen bei $r = .395$. Die mittlere Korrelation über alle Bewertungen betrug $r = .336$. Es waren keine systematischen Unterschiede zwischen bestimmten Kriterien oder zwischen den unmittelbaren und getrennten Wiederholungen zu erkennen. In der einschlägigen Literatur zum Lenkgefühl werden keine entsprechenden Retest-Reliabilitäten berichtet, mit denen diese Ergebnisse verglichen werden könnten. In der psychologischen Forschung zu Persönlichkeitstests werden durchschnittliche Retest-Reliabilität von .79 berichtet (Schmidt-Atzert & Amelang, 2012). Die Vergleichbarkeit mit diesem Wert ist jedoch nicht nur wegen der verschiedenen inhaltlichen Ausrichtung eingeschränkt. Diese Korrelation bezieht sich auf den Zusammenhang zwischen Testwerten, welche aus allen Items eines Fragebogens berechnet wurden. Da den Korrelationen im vorliegenden Fall jedoch einzelne Items zugrunde liegen, ist von insgesamt schwächeren Zusammenhängen auszugehen. Es bleibt somit offen, wie die berechneten Retest-Reliabilitäten einzuordnen sind. Neben der Retest-Reliabilität wurden auch die internen Konsistenzen bestimmt (Lienert & Raatz, 1998). Hierfür wurde Cronbachs α herangezogen (Cronbach, 1951). Dieser Wert entspricht der gemittelten

Testhalbierungsreliabilität eines Tests für alle möglichen Testhalbierungen (Bortz & Döring, 2006). Für die beiden Faktoren lag Cronbachs α bei .705 und .595. Diese Werte ließen in Abhängigkeit der Itemanzahlen und mittleren Korrelation (Cortina, 1993) den Schluss zu, dass eine ausreichende interne Konsistenz vorliegt. Die gefundenen Werte sind jedoch niedriger als das Cronbachs α von .9, welches für einen Fragebogen zur Lenkungsbewertung mit sieben Items gefunden wurde (Riedel & Arbinger, 1997). Da die zugrunde liegenden Daten aus einem Versuch mit verschiedenen Fahrzeugen gewonnen wurden, könnte jedoch auch hier die allgemein niedrigere Varianz zu einer höheren internen Konsistenz geführt haben. Ferner lag der Skala eine größere Anzahl an Items zugrunde, was zu einer Erhöhung von Cronbachs α führt (Field, 2013). Insgesamt lassen die Ergebnisse den Schluss zu, dass für das Versuchskonzept eine ausreichende Reliabilität vorliegt.

Abschließend wurde die Validität betrachtet, welche das Ausmaß beschreibt, zu dem ein Test das misst, was er messen soll (Renner et al., 2012). Die Inhaltsvalidität gibt an, inwiefern die Items eines Fragebogens das zu messende Merkmal abdecken (Moosbrugger & Kelava, 2012). Im Rahmen dieser Arbeit wurde sie durch das Offenlegen der Vorgehensweise bei der Fragebogenentwicklung belegt. Die Überprüfung der Kriteriumsvalidität gestaltet sich oftmals schwierig. Im vorliegenden Fall konnten keine Kriterien außerhalb der Testsituation gefunden werden, welche sinnvoll mit den subjektiven Bewertungen in Zusammenhang gebracht werden konnten. Dieselbe Problematik trat auch für die Konstruktvalidität auf. Somit konnte bezüglich der Messgenauigkeit nur die Inhaltsvalidität untersucht werden. Zusammenfassend wurden keine Einschränkungen festgestellt. Bei nachfolgenden Studien sollte jedoch stärkeres Augenmerk darauf gelegt werden, eine detailliertere Analyse der Validität zu ermöglichen.

Alles in allem wurde der Fragebogen mit einer Faktoren-, Item- und Reliabilitätsanalyse untersucht. Außerdem wurden die Gütekriterien Objektivität, Reliabilität und Validität beleuchtet. Insgesamt legten die Ergebnisse nahe, dass der Fragebogen eine ausreichende psychometrische Qualität aufwies. Hypothese 1 wurde beibehalten.

4.4.2. Vorhersagbarkeit der Gefallen-Bewertungen in Abhängigkeit der Niveau-Bewertung

H2: Die Gefallen-Bewertungen lassen sich in Abhängigkeit von den Niveau-Bewertungen vorhersagen.

Die in diesem Versuch verwendete Skala basiert auf der Annahme, dass die subjektive Bewertung in eine Niveau- und Gefallen-Bewertung aufgeteilt werden kann. Inwiefern das individuelle Gefallen auf Basis der quasi-objektiven Einschätzung vorhergesagt werden kann, wurde in dieser Hypothese überprüft. Für alle Varianten konnten Regressionsmodelle berechnet werden, welche die Gefallen-Bewertungen signifikant vorhersagten. Bei fünf der Variablen lag ein quadratischer Zusammenhang zwischen der Niveau- und der Gefallen-Bewertung vor. Für die Kriterien *Rückstellverhalten*, *Lenkpräzision* und *Mittenzentrierung* zeigte die Hinzunahme des quadratischen Faktors jedoch keine signifikante Verbesserung, weshalb hier das lineare Modell angenommen wurde. Die Hypothese wurde beibehalten.

Im Folgenden soll genauer auf die zugrunde liegenden Bewertungen eingegangen werden, allen voran auf die Varianzaufklärung. Beim *Anlenkmoment* war diese mit 15 % am niedrigsten. Auch die Gefallen-Bewertungen von *Lenkmoment (Slalom)* konnten nur zu 18 % vorhergesagt werden. Mittlere Varianzaufklärungen lagen bei den Kriterien *Haltemoment* (37 %), *Mittenzentrierung* (54 %) und *Lenkmoment (Rangieren)* (58 %) vor. Die

höchsten Varianzaufklärungen traten bei den Kriterien *Fahrbahnkontakt* (78 %), *Lenkpräzision* (83 %) und *Rückstellverhalten* (84 %) auf. Dies entspricht einer mittleren Varianzaufklärung von 53 %.

Diese Varianzaufklärungen werden nun in Bezug zu anderen Forschungsergebnissen gebracht. Decker (2009) berichtet Korrelationskoeffizienten zwischen quasi-objektiven und subjektiven Bewertungen. Die geteilte Varianz war mit 10 % beim Gierverhalten am niedrigsten. Dieser Wert liegt noch unterhalb der niedrigsten, in diesem Versuch gefundenen Varianzaufklärungen. Im Gegensatz hierzu lag mit 86 % geteilter Varianz bei dem Kriterium Überschwinger Gieren ein enger Zusammenhang vor. Für den Lenkmomentverlauf lag die geteilte Varianz bei 72 %, was oberhalb der hier gefundenen Varianzaufklärung für die Kriterien zum Lenkmoment lag. Bei Zschocke (2009) finden sich Ergebnisse aus Regressionen, die jedoch auf Bewertungen durch Experten beruhen. Es wurden lineare Zusammenhänge zwischen Niveau- und Gefallen-Bewertungen mit geteilten Varianzen von über 90 % gefunden. Diese traten beispielsweise für die Bewertungen des Parkiermoments, der Lenkpräzision oder des Geradeauslaufs auf. Andererseits wurden für die Bewertungen der Mittenzentrierung, des Anlenkmoments und des Haltemoments deutlich schwächere Zusammenhänge gefunden. Da jedoch lediglich angegeben wurde, dass die geteilte Varianz unter 50 % lag, kann kein genauerer Vergleich erfolgen. Somit lag auch in diesen Studien eine große Variabilität bzgl. der Größe der aufgeklärten Varianzen vor. Die Größe der gefundenen Varianzaufklärungen war ebenfalls vergleichbar.

Im Anschluss sollen die Regressionskurven selbst beleuchtet werden. Für *Lenkmoment (Rangieren)* zeigte sich, dass das Gefallen mit steigender Niveau-Bewertung beständig abnahm. Für *Lenkmoment (Slalom)* lag ein ausgeprägter negativer quadratischer Zusammenhang vor, welcher für eine mittlere Niveau-Bewertung die höchste Gefallen-Bewertung prädizierte. Auch für das *Anlenkmoment* lag ein negativer, quadratischer Zusammenhang vor. Diese Regressionskurve war jedoch deutlich flacher. Zusammen mit der niedrigen Varianzaufklärung lässt sich daraus schließen, dass den Bewertungen keine über alle Probanden hinweg bestehende Präferenz zugrunde liegt. Die Kurve des *Haltemoments* hatte ihr Maximum bei niedrigen Niveau-Bewertungen. Für die übrigen vier Kriterien *Rückstellverhalten*, *Lenkpräzision*, *Fahrbahnkontakt* und *Mittenzentrierung* gingen höhere Niveau-Bewertungen mit höheren Gefallen-Bewertungen einher.

Die graphische Analyse der Punktdiagramme zeigt bei den Kriterien *Lenkmoment (Slalom)*, *Anlenkmoment* und *Haltemoment* allerdings auch Limitationen der Vorhersage auf. So ist bei diesen Variablen neben den Werten im Bereich der Regressionsgeraden eine Diagonale mit $y \sim -x$ erkennbar. Diesen Bewertungen liegen offensichtlich andere Mechanismen zugrunde. Anhand der Varianzaufklärung lässt sich dann erkennen, wie stark die zugrunde liegende Diskrepanz zwischen den individuellen Präferenzen ausgeprägt ist. Dies entspricht der dem Bewertungssystem zugrunde liegenden Annahme, dass dieselbe quasi-objektive Niveau-Einschätzung zu unterschiedlichen individuellen Gefallen-Bewertungen führen kann (Decker, 2009; Zschocke, 2009). Da diese Bewertungskriterien mit ihrem Fokus auf das aufzuwendende Lenkmoment den Lenkcharakter eines Fahrzeugs beschreiben, ist die Individualität der Präferenzen in einem besonderen Maße ausgeprägt (Zschocke, 2009). Dies spiegelte sich letztendlich auch in den niedrigen Varianzaufklärungen wider.

Für die übrigen Kriterien *Rückstellverhalten*, *Lenkpräzision*, *Fahrbahnkontakt* und *Mittenzentrierung* ging ein Zuwachs in der Ausprägung mit nahezu demselben Zuwachs des Gefallens einher. Dies ist mit der inhaltlichen Natur der Kriterien begründbar. Da beim *Rückstellverhalten* bewertet wird, inwiefern die Variante das Zurückstellen unterstützt, ist hier ein positiv-linearer Zusammenhang zwischen Ausprägung und Gefallen naheliegend. Der gleiche Mechanismus ist auch für die übrigen drei Kriterien anzunehmen. Somit ist hier die Individualität des Zusammenhangs deutlich geringer, was sich entsprechend in größeren Varianzaufklärungen und der besseren Vorhersage der Gefallen-Bewertung zeigte. Dies gilt zumindest in dem Rahmen, der mit den Varianten abgedeckt wurde. So könnte beispielsweise ein noch stärker ausgeprägtes Rückstellverhalten zu niedrigeren Gefallen-Bewertungen führen, da es beispielsweise als zu stark empfunden wird oder zum Überlenken über die Geradeausstellung kommt.

Alles in allem konnten also die Gefallen-Bewertung aller Kriterien auf Basis der Niveau-Bewertung vorhergesagt werden. Aus den teilweise geringen Varianzaufklärungen lässt sich für nachfolgende Studien ableiten, dass das zweistufige Vorgehen der Bewertung von Niveau und Gefallen beibehalten werden sollte.

4.4.3. Zuverlässigkeit der Bewertung

H3: Die wiederholte Bewertung der gleichen Variante der Lenkunterstützung führt nicht zu unterschiedlichen Bewertungen.

Diese Hypothese diente zur Überprüfung, wie zuverlässig die Bewertungen der Normalfahrer sind. Deshalb sah das Versuchsdesign die Wiederholung von Varianten vor, um dann die einzelnen Bewertungen miteinander vergleichen zu können. Beim Ausbleiben von Unterschieden waren die Bewertungen als zuverlässig einzustufen.

H3a: Die wiederholte Bewertung der gleichen Variante führt nicht zu unterschiedlichen Niveau- und Gefallen-Bewertungen (Re-Test).

In der ersten Unterhypothese wurde überprüft, ob sich die erneute Bewertung der gleichen Variante von der ursprünglichen Bewertung unterscheidet. Bei 44 von 46 Vergleichen lagen gemäß der Hypothese keine signifikanten Unterschiede vor, weshalb die Hypothese beibehalten wurde. In den zwei Fällen lagen signifikante Unterschiede mit großen Effektstärken vor. In beiden Fällen wurden Varianten mit höheren aufzuwendenden Lenkmomenten evaluiert, wobei die erneute Bewertung signifikant höher ausfiel. Hinsichtlich der Art der Wiederholung ergab sich jedoch keine Regelmäßigkeit, da in einem Fall eine unmittelbare, im anderen eine getrennte Wiederholung vorlag.

Ein signifikanter Unterschied lag beim Kriterium *Haltemoment* vor, welches in Kapitel 4.4.4 tiefer diskutiert wird. Der andere Unterschied wurde bei der Bewertung der *Mittenzentrierung* gefunden. Die Versuchsleiter dokumentierten bereits während der Versuche Verständnisprobleme bei der Bewertung dieses Kriteriums. Aus den verbalen Anmerkungen bei der Bewertung ging hervor, dass einige Probanden das Kriterium trotz der Erklärung und entsprechenden Hinweisen als das Ausbleiben von Lenkungsspiel interpretierten. Obwohl die subjektive Bewertung der beiden Aspekte gewissermaßen zusammenhängt, wurde dennoch nicht der gewünschte Aspekt bewertet. Andere Probanden entwickelten erst im Laufe der sechs Varianten ein Verständnis für die Bewertung des Mittengefühls, weshalb es auch deshalb zu unzuverlässigen Bewertungen gekommen sein könnte.

Alles in allem könnte also die Schwierigkeit bei der Bewertung des Kriteriums zu den signifikant verschiedenen Bewertungen geführt haben.

Aufgrund der teilweise kleinen Stichprobengröße wurde eine Analyse der nicht signifikanten Vergleiche angeschlossen. Es zeigte sich für die fünf Vergleiche mit mittleren Effektstärken kein Muster, dass stets die ursprüngliche oder erneute Bewertung höher ausfiel. Vielmehr fiel auf, dass hauptsächlich die indirekt mit dem Lenkmoment zusammenhängenden Kriterien betroffen waren. Allen voran ist hier die *Lenkpräzision* zu nennen, bei der drei der fünf mittleren Effektstärken auftraten. Ferner lag in allen Fällen eine getrennte Wiederholung vor. Aus diesen Punkten lässt sich trotz der nicht-signifikanten Vergleiche möglicherweise ableiten, dass die Fahrer Schwierigkeiten hatten, diese indirekten Kriterien bei einer getrennten Wiederholung zuverlässig zu bewerten. Dies könnte zum einen, wie bei der *Mittenzentrierung* bereits diskutiert, in der geringeren Verständlichkeit oder Konkretheit der Kriterien begründet sein. Bei der *Lenkpräzision* könnte dies noch stärker der Fall sein, da sie sich aus vielen verschiedenen Facetten zusammensetzt. Als Einflussfaktoren sind beispielsweise Reibung, Dämpfung und Zeitverzug zu nennen (Koch, 2010). Diese Faktoren könnten allgemein zu einer größeren Varianz in den Bewertungen geführt haben, was aufgrund des Abstands zwischen den identischen Varianten möglicherweise besonders bei der getrennten Wiederholung zum Tragen kam.

Dennoch ist hervorzuheben, dass bei einem Großteil der getrennten Wiederholungen keine unterschiedlichen Bewertungen existierten. Dies entspricht auch den Ergebnissen von Studie II (Kapitel 3.3.3), bei der die wiederholte Bewertung innerhalb des Versuchs insgesamt nicht zu unterschiedlichen Werten führte. Im Gegensatz zu diesem Versuch konnten die Probanden nun die verschiedenen Varianten viel unmittelbarer vergleichen, als ihnen dies mit den verschiedenen Fahrzeugen möglich war. Dieses Ergebnis spricht aus mehreren Gründen besonders für die Eignung der Normalfahrer als Beurteiler. Zum einen war die Wahrscheinlichkeit einer zufällig ähnlichen Bewertung relativ gering. Schließlich wurde den Probanden mit sechs Varianten eine große Anzahl an Alternativen angeboten. Weiterhin war auch die Spanne zwischen den Varianten hinsichtlich der objektiven Parameter größer als in den vorherigen Versuchen, was sich in einer Verteilung der Beurteilungen über die gesamte Skala zeigte. Außerdem war auch der Großteil der Bewertung der getrennten Wiederholungen nicht verschieden. Obwohl also zwischen diesen beiden Messzeitpunkten vier andere Varianten gefahren und bewertet wurden, blieben die Bewertungen konstant.

Die Ergebnisse decken sich auch mit anderen Studien, wonach eine von Normalfahrern zu Beginn und Ende des Sets evaluierte Variante nicht zu verschiedenen Bewertungen führte Anand (2014).

H3b: Die Niveau- und Gefallen-Bewertungen zeigen keinen Unterschied zwischen der Referenz und der blind als Variante gefahrenen Referenz an.

Im zweiten Teil der Hypothese wurden die Bewertungen der als Variante getarnten Referenz analysiert. Die Rohwerte wurden gegen den Wert 0 getestet, da dieser auf der Skala für keinen wahrgenommenen Unterschied stand. Lediglich in 1 von 16 Fällen war die Bewertung der Variante signifikant höher als 0. In einem weiteren Fall war die Bewertung vor der α -Fehler-Korrektur signifikant höher. Es lagen in beiden Fällen kleine Effektstärken vor. Somit wurde die Hypothese beibehalten.

Dennoch soll auf die signifikanten Ergebnisse eingegangen werden. Diese lagen den Bewertungen des Kriteriums *Rückstellverhalten* zugrunde, welches genauso wie das in den Analysen zu H3a auffällige Kriterium *Haltemoment* im Oval evaluiert wurde. Die größere Varianz in der subjektiven Bewertung dieser Kriterien könnte möglicherweise durch die Ausführung des Fahrmanövers verursacht worden sein. Bei diesem Manöver war es aufgrund des Ausschwenkens des Aufliegers nicht möglich, die Fahrspur wie bei den anderen Manövern eng mit Pylonen vorzugeben. Das daraus resultierende variabelere Lenkverhalten könnte zu einer weniger zuverlässigen Bewertung geführt haben. Bei der Bewertung des *Rückstellverhaltens* kann dies sogar in einem noch stärkeren Maße zutreffen, da der Kurveneingangswinkel und die Art und Weise, wie die Kurve angefahren wird, das Rückstellverhalten maßgeblich beeinflussen. Ferner zeigten die Probanden in einigen Durchgängen entgegen der expliziten Aufforderung des Versuchsleiters teilweise ein realitätsfernes Verhalten. Anstatt das Lenkrad entsprechend mitzuführen, ließen es einige Probanden völlig los und beobachteten dessen Rücklauf lediglich.

Insgesamt ist also festzuhalten, dass die Bewertung der als Variante getarnten Referenz mit Ausnahme des *Rückstellverhaltens* nicht zu signifikant verschiedenen Bewertungen führte. Dies muss insbesondere vor dem Hintergrund gesehen werden, dass die Fahrer aufgrund des Blindversuchs von sechs tatsächlich unterschiedlichen Varianten ausgingen. Im Sinne der sozialen Erwünschtheit (Turner & Martin, 1984) könnte es das Versuchsdesign also gefördert haben, dass die Fahrer eher einen Unterschied berichten. Aufgrund der vorgestellten Ergebnisse zu den Hypothesen 3a und 3b wurde die Hypothese 3 beibehalten. Dennoch sollten die aufgedeckten Unterschiede zum *Rückstellverhalten* zu einem späteren Zeitpunkt genauer beleuchtet werden.

4.4.4. Unterschiedliche Niveau-Bewertung der Varianten

H4: Die objektiv vorhandenen Unterschiede in der Höhe des aufzuwendenden Lenkmoments spiegeln sich in entsprechend unterschiedlichen Niveau-Bewertungen wider.

Zur Überprüfung dieser Annahme wurden ANOVAs mit Messwiederholung berechnet. Mit diesen wurden die Niveau-Bewertungen der Varianten in den vier Kriterien *Lenkmoment (Rangieren)*, *Lenkmoment (Slalom)*, *Anlenkmoment* und *Haltemoment* analysiert. Alle Haupteffekte waren höchst signifikant. Die angeschlossenen wiederholten Kontraste zeigten für alle Kriterien signifikante Unterschiede zwischen den Bewertungen V1 und V2, V2 und V3, V3 und V4 sowie V4 und V5. Den Unterschieden lagen große Effektstärken zugrunde. Die Hypothese wurde deshalb beibehalten.

Zusätzlich wurden prozentuale Häufigkeiten ausgezählt. Den Unterschied von 1 Nm nahmen 79 % der Fahrer wahr. Der Unterschied von 0,7 Nm schlug sich bei 75 % bzw. 83 % in unterschiedlichen Bewertungen nieder. Schlussendlich wurde die kleinste Differenz von 0,4 Nm im maximalen Lenkmoment von 65 % der Fahrer entdeckt. Somit waren die Prozentsätze an Fahrern, die die Unterschiede von 0,7 Nm bzw. 1,0 Nm wahrnahmen, vergleichbar groß. Bei dem Unterschied von 0,4 Nm war der Anteil der entdeckten Unterschiede jedoch bis auf eine Ausnahme bei allen Bewertungskriterien kleiner als bei 0,7 Nm bzw. 1,0 Nm.

Allgemein sind die Ergebnisse vergleichbar mit denen aus anderen Studien zu just-noticeable-differences (Buschardt, 2003; Hiraoka et al., 2008; Newberry et al., 2007; Wolf, 2009). Es muss jedoch auf zwei Unterschiede in der Methode hingewiesen werden: In den vorherigen Studien lag der Fokus der Untersuchungen explizit darauf,

Wahrnehmungsschwellen zu berechnen. Deshalb fanden diese auch im Labor an entsprechenden Prüfständen statt. Im vorliegenden Fall jedoch lag der Fokus auf der Erprobung verschiedener Varianten mit dem Fokus auf ihre subjektive Bewertung. Deshalb fand die Bewertung während der Fahrt statt. Aufgrund der unterschiedlichen Zielsetzung und dem daraus abgeleiteten Versuchsdesign konnten bei diesem Versuch nur die V2 und V4 mit V3 als Referenz unmittelbar miteinander verglichen werden. Dies war dem geschuldet, dass der Versuchsablauf auf einen Vergleich der Varianten mit der Referenz ausgerichtet war, jedoch nicht auf einen direkten Vergleich der Varianten untereinander.

Diese realitätsnahe Erprobung könnte jedoch auch einen Einfluss auf die Wahrnehmung gehabt haben. Insbesondere beim Kriterium *Haltemoment* lagen widersprüchliche Ergebnisse vor. Einerseits war der Anteil an Fahrern, die den Unterschied zwischen V2 und V3 entdeckten, hier am größten. Andererseits wurde jedoch zwischen V4 und V5 der niedrigste Anteil an Fahrern gefunden, die den zugrunde liegenden Unterschied korrekt wahrnahmen. Gleichzeitig war der Anteil an Fahrern, die den Unterschied entgegengesetzt wahrnahmen, mit knapp 30 % am höchsten. Möglicherweise könnten die unterschiedlichen Ergebnisse im Manöver begründet sein. Das Oval, in dem das *Haltemoment* erprobt wurde, konnte nicht so eng vorgegeben werden wie die anderen Manöver (Kapitel 4.4.3). Deshalb könnten die gewählten Lenkwinkel sowie die Dauer, für die das Lenkrad in der eingelenkten Position gehalten wurde, variiert haben. Diese Unterschiede im Fahrverhalten könnten zu unterschiedlichen Bewertungen geführt haben.

Alles in allem zeigten die Daten, dass die Fahrer Unterschiede von 0,4 Nm, 0,7 Nm und 1,0 Nm im maximalen Lenkmoment im Fahrbetrieb wahrnehmen konnten. Dies ist einerseits dahingehend bedeutend, als dass nun mehr Ergebnisse diesbezüglich aus dem schweren Nutzfahrzeug vorliegen. Zum anderen handelt es sich hierbei um eine entscheidende Basis für die weiteren Analysen der Ergebnisse.

4.4.5. Vorhersagbarkeit der Niveau-Bewertung in Abhängigkeit der Variante

H5: Die Niveau-Bewertungen der direkt vom aufzuwendenden Lenkmoment abhängigen Bewertungskriterien lassen sich in Abhängigkeit von Parametern der Lenkungsauslegung vorhersagen.

Die Ergebnisse der berechneten Regressionen unterstützten die Hypothese, dass die Niveau-Bewertungen aus den objektiven Kennwerten vorhergesagt werden können. Als Prädiktor wurde das maximal vom Fahrer aufzubringende Lenkmoment der Varianten herangezogen. In allen Fällen lagen quadratische Zusammenhänge vor. Bei den vier Kriterien *Lenkmoment (Rangieren)*, *Lenkmoment (Slalom)*, *Anlenkmoment* und *Haltemoment* gingen, wie bei H4 bereits überprüft, höhere aufzuwendende Lenkmomente mit entsprechend höheren Niveau-Bewertungen einher. Die Niveau-Bewertungen des *Rückstellverhaltens* waren bei Varianten mit mittleren aufzuwendenden Lenkmomenten am höchsten. Für diese Varianten lagen auch objektiv messbar die größten Rückstellmomente vor.

Die Varianzaufklärung betrug bei der Bewertung des Kriteriums *Lenkmoment (Rangieren)* 77 %, beim *Haltemoment* 79 %, beim *Anlenkmoment* 80 % und beim *Lenkmoment (Slalom)* 83 %. Somit lag bei diesen Kriterien ein enger Zusammenhang zwischen dem vom Fahrer aufzuwendenden Lenkmoment und den Niveau-Bewertungen vor. Dies spiegelt sich auch darin wieder, dass die Streuung der Bewertungen einer Variante

vergleichsweise klein war. Im Gegensatz dazu stehen die Niveau-Bewertungen des *Rückstellverhaltens*, welche nur zu 9 % vorhergesagt werden konnten. Dies spiegelte sich auch in den größeren Streuungen innerhalb der Varianten wider.

Die gefundenen Varianzaufklärungen sollen nun in Bezug zu anderen Forschungsergebnissen gebracht werden. Es wird an dieser Stelle jedoch nur auf solche eingegangen, die sich auf quasi-objektive Bewertungen beziehen (siehe Unterscheidung in Kapitel 2.3.3). Im Gegensatz zu dem vorliegenden Vorgehen liegen in den meisten Fällen multiple Regressionsmodelle zugrunde.

Zuerst wird auf die subjektive Bewertung des Lenkmoments eingegangen. Bei Zschocke (2009) betrug die Varianzaufklärung der subjektiven Bewertung des Anlenkmoments auf Basis des objektiv messbaren Anlenkmoments 76 %. In einer Fahrsimulator-Studie konnte die subjektive Bewertung des Lenkmoments sogar zu 93 % aus dem Lenkmoment und dem Lenkmomentgradienten vorhergesagt werden (Dang et al., 2014). Auf Basis des maximalen Lenkmoments und des Lenkwinkels betrug die Varianzaufklärung des subjektiv wahrgenommenen Lenkmoments in einer Studie am Lkw 82 % (Rothhämel, 2013). Für die quasi-objektive Bewertung des Lenkmomentverlaufs lag unter Nutzung von open-loop-Kennwerten als Prädiktor eine geteilte Varianz von 61 % vor (Decker, 2009). Bei Kennwerten, die für jeden Fahrer individuell aus den Manövern berechnet wurden, lag dieser Anteil bei 46 %.

Bezüglich des Rücklaufs liegen teilweise ebenfalls sehr gute Vorhersagen von bis zu 95 % vor (Dang et al., 2014), wenn das Regressionsmodell aus Lenkmoment, Lenkmomentgradienten und Steifigkeit besteht. Bei Rothhämel (2013) konnten über 70 % Varianz im subjektiven Rückstellverhalten vorhergesagt werden, wenn Querbeschleunigung, benötigter Lenkwinkel sowie Lenkmoment als Prädiktoren fungierten. Aus Expertenbeurteilungen, die sich nur auf einen Prädiktor stützen, liegt eine Varianzaufklärung von 50 % für den Rücklauf vor (Zschocke, 2009).

Insgesamt ergibt sich aus allen Studien das Bild, dass die erzielte Varianzaufklärung beim Lenkmoment höher ist als beim Rückstellverhalten. Jedoch waren die Vorhersagen des subjektiven Rückstellverhaltens in anderen Studien deutlich besser als die, die auf Basis der Daten dieser Studie möglich waren. Eine Erklärung könnte erneut die größere Variabilität sein, mit der das zur Bewertung genutzte Fahrmanöver durchfahren wurde. Es ist jedoch wahrscheinlicher, dass der Prädiktor nicht gut geeignet ist zur Vorhersage des Rückstellverhaltens.

4.4.6. Vorhersagbarkeit der Gefallen-Bewertungen in Abhängigkeit der Variante

H6: Die Gefallen-Bewertungen der direkt vom aufzuwendenden Lenkmoment abhängigen Bewertungskriterien lassen sich in Abhängigkeit von Parametern der Lenkungsauslegung vorhersagen.

Die Ergebnisse der berechneten Regressionen unterstützten die Hypothese, dass die Gefallen-Bewertungen aus den objektiven Kennwerten vorhergesagt werden können. Als Prädiktor wurde das maximal vom Fahrer aufzubringende Lenkmoment der Varianten herangezogen. Es zeigten sich in allen fünf Fällen negative quadratische Zusammenhänge. Die Varianzaufklärung betrug bei der Bewertung von *Lenkmoment (Rangieren)* 45 %, was für eine vergleichsweise gute Vorhersage steht. Die Varianzaufklärung war bei den Kriterien *Lenkmoment (Slalom)* und *Anlenkmoment* mit 14 % und 18 % deutlich niedriger. Für das *Haltemoment* war mit einer Varianzaufklärung

von 29 % ebenfalls eine entsprechend gute Vorhersage möglich. Das *Rückstellverhalten* konnte zu 9 % vorhergesagt werden. Für das Kriterium *Lenkmoment (Rangieren)* lag das Maximum der Gefallen-Bewertung im unteren Bereich des aufzuwendenden Lenkmoments. Für alle übrigen Kriterien lagen die Maxima jeweils im mittleren Bereich des benötigten Lenkmoments.

Die Varianzaufklärungen aller fünf Kriterien waren durchweg niedriger als die bei der Vorhersage der Niveau-Bewertung. Während sich diese teilweise zu über 80 % vorhersagen ließen, konnte hier nur weniger als die Hälfte der Varianz erklärt werden. Dies ist auch beim Vergleich der Regressionsdiagramme ersichtlich. Während beim Niveau eine deutliche Häufung der Bewertungen entlang der Regressionsgerade ersichtlich ist, verteilen sich die Bewertungen des Gefallens über den gesamten Skalenraum. Einzige Ausnahme war das *Rückstellverhalten*, bei welchem auch die Niveau-Bewertungen große Streuungen aufwiesen. Aus den großen Streuungen wird, wie bereits in Kapitel 4.4.2 erläutert, die Individualität der Gefallen-Bewertung deutlich. Die etwas höheren Varianzaufklärungen bei *Lenkmoment (Rangieren)* und *Haltemoment* legen nahe, dass die Präferenzen bezüglich des Lenkmoments bei langsamer Fahrt noch vergleichsweise ähnlich waren. Die hohen Varianzaufklärungen dieser beiden Kriterien stehen jedoch auch im Widerspruch zu den Ergebnissen zu H3 (Kapitel 4.4.3), bei denen sie bzgl. der Zuverlässigkeit auffällig waren. Es bleibt jedoch offen, warum die Varianzaufklärung bei *Lenkmoment (Slalom)* derart niedrig war. Möglicherweise sind die individuellen Präferenzen in diesem mittleren Geschwindigkeitsbereich besonders unterschiedlich.

Insgesamt entsprechen die Ergebnisse aber auch denen anderer Studien. Hier konnte bspw. die subjektive Bewertung des Lenkmomentverlaufs zu 42 % aus open-loop-Kennwerten vorhergesagt werden (Decker, 2009). Wurde der Kennwert für jeden Fahrer individuell aus den Manövern berechnet, lag dieser Anteil bei 35 %.

4.4.7. Explorative Analysen

EXP: Unterscheidet sich die Zuverlässigkeit der Bewertungen in Abhängigkeit der Gesamtfahrerfahrung und des Alters?

Die Fahrer wurden entsprechend ihres Alters und ihrer Gesamtfahrerfahrung in jeweils drei Gruppen aufgeteilt und anschließend die Analysen zu H3b wiederholt.

Die Analysen des Alters zeigten, dass die Bewertungen der Fahrer mit mittlerer und hoher Fahrerfahrung signifikant größer als der Wert 0 waren. Das bedeutet, dass sie die als Variante getarnte Referenz höher bewerteten als die Referenz selbst. Dieses Ergebnis steht im Widerspruch zu den Annahmen des Versuchskonzepts. Wie eingangs angemerkt, wurde eine Mindest-Fahrerfahrung für Probanden gefordert (Ackert, 2008; Ellmann, 2003; Fastenmeier, 2005). So sollte sichergestellt werden, dass sich die Probanden neben der Fahraufgabe auf die Bewertung des Lenkgefühls konzentrieren können. Gleichzeitig war mit einer entsprechenden Fahrerfahrung von zuverlässigeren Bewertungen auszugehen. Dies kann u. a. damit begründet, dass erfahrene Fahrer ein flüssigeres Lenk- und Fahrverhalten aufweisen (Abendroth & Bruder, 2015) und somit die Manöver über mehrere Durchgänge hinweg konstant durchfahren. Im vorliegenden Fall entsprachen jedoch nur die Bewertungen der weniger erfahrenen Probanden den tatsächlichen Begebenheiten.

Bei der Analyse des Alters zeigte sich, dass die jüngeren Fahrer die Referenz höher einschätzten, wenn sie sie als Variante bewerteten. Dies könnte nur darin begründet werden, dass die jüngeren Fahrer aufgrund ihrer noch geringen Fahrpraxis weniger zuverlässig bewerteten. Diese Begründung stünde allerdings im Gegensatz zum vorherigen Ergebnis. Es liegt somit eine gewisse Gegensätzlichkeit der Ergebnisse zum Fahrverhalten und zum Alter vor. Die Vermutung, dass eine höhere Gesamtfahrfahrung häufig mit einem höheren Alter einhergeht, konnte auch statistisch mit einer signifikanten Korrelation zwischen den beiden Variablen bestätigt werden ($r = .241, p = .040$).

Insgesamt bleibt zu bedenken, dass die Effektstärken aller drei signifikanten Vergleiche unter der Grenze für einen kleinen Effekt lagen. Diese Unterschiede waren also als deutlich kleiner als die sonst bei dieser Studie vorkommenden Effektstärken. Da die Ergebnisse gegenläufig waren, war von keiner Systematik auszugehen. Die Anforderungen an die Stichprobe wurden deswegen unverändert beibehalten.

4.4.8. Ableitung der Gestaltungsempfehlungen

Für jedes Manöver wurden die entsprechend der Gefallen-Bewertungen favorisierten Varianten ermittelt. Die Rangiersituation wurde mit einer Fahrgeschwindigkeit von max. 5 km/h absolviert. Bezüglich des *Lenkmoments* (*Rangieren*) wurden V1 und V2 mit den geringsten aufzuwendenden Lenkmomenten bevorzugt. Es wurde somit keine Variante angeboten, die in dieser Situation subjektiv als zu leichtgängig eingeschätzt wurde. Es kann deshalb nicht ausgeschlossen werden, dass ein noch niedrigeres aufzuwendendes Lenkmoment zu noch besseren Gefallen-Bewertungen geführt hätte. Für das Oval (15 km/h), welches Kreisverkehrsdurchfahrten oder Abbiegesituationen im städtischen Bereich abbildete, wurden V1 und V2 hinsichtlich des *Haltemoments* bevorzugt. Das ebenfalls im Oval bewertete *Rückstellverhalten* wurde für V2 und V3 am besten eingeschätzt. Für den Slalom mit 45 km/h waren V2 und V3 Favoriten hinsichtlich des *Lenkmoments* (*Slalom*). Hinsichtlich des ebenso in diesem Manöver bewerteten *Fahrbahnkontakts* gefielen die Varianten V3, V4 und V5 ähnlich gut. Beim doppelten Spurwechsel (45 km/h) wurden V2 und V3 bezüglich des *Anlenkmoments* favorisiert. Bei der *Lenkpräzision* fiel nur V1 mit einer deutlich schlechteren Bewertung als alle anderen Varianten auf. Bei der Geradeausfahrt wurde die *Mittenzentrierung* der V3, V4 und V5 am besten bewertet.

Insgesamt konnten für jede Fahrsituation geeignete Varianten angeboten werden. In den Bewertungen spiegelt sich außerdem der Einfluss der Geschwindigkeit auf die Präferenzen wider. Wie in der Literatur zu finden (Bertolini & Hogan, 1999; Kim & Song, 2002), werden bei höheren Geschwindigkeiten höhere aufzuwendende Lenkmomente bevorzugt.

Dennoch sind zwei Punkte anzumerken: Einerseits schienen die Probanden bei den Kriterien *Fahrbahnkontakt*, *Lenkpräzision* und *Mittenzentrierung* hinsichtlich der favorisierten Variante nicht stark zu differenzieren. In diesen Fällen wurden drei bis vier der fünf Varianten ähnlich bewertet. Jedoch zeigten die Analysen zu H4 (Kapitel 4.3.4), dass die Probanden die den Varianten zugrunde liegenden Unterschiede im Lenkmoment wahrnahmen. Hieraus könnte geschlossen werden, dass die Varianten keinen großen Einfluss auf die Bewertung dieser Kriterien haben. Erst bei extrem niedrigen Lenkmomenten könnte ein Lenkgefühl einsetzen, was von Normalfahrern oftmals als schwimmend oder schwammig bezeichnet wird.

Ferner war festzustellen, dass die Standardabweichungen der Bewertungen zu V1 insgesamt größer waren als die der anderen Kriterien. Die große Standardabweichung bedeutet, dass die Fahrer sehr unterschiedliche, individuelle Präferenzen aufweisen. Dies entspricht den Erkenntnissen aus Studie II, in der die Fahrer das Fahrzeug mit der hybriden Lenkung ebenfalls sehr heterogen hinsichtlich des Gefallens bewerteten. Somit gibt es einige Fahrer, die eine eher schwergängige Lenkungsauslegung wie bei hydraulisch unterstützten Lkw bevorzugen. Andere Fahrer wiederum bevorzugen sehr niedrige Lenkmomente wie sie im Pkw-Bereich zu finden sind. Generell geht das Ergebnis mit anderen Studien einher, dass Fahrer extreme und ihnen unbekannte Lenkmomentgestaltungen hinsichtlich Komfort und Kontrollierbarkeit niedrig bewerten (Anand, 2014). Der Individualität der Vorlieben kann dadurch begegnet werden, dass der Fahrer selbst die Möglichkeit erhält, die Einstellungen an seine Bedürfnisse anzupassen (Barthenheier, 2004). Dies wird heute mithilfe von Fahrprogrammen realisiert, welche jedoch nur einen sehr kleinen Verstellbereich zulassen (Anand et al., 2011).

4.4.9. Limitationen des Versuchs

Versuchsaufbau

Insgesamt erwies sich das finale Versuchskonzept mit der relativen Bewertung der Varianten in Bezug auf eine Referenz als zielführend. In den vorherigen Versuchen hatten die Probanden Schwierigkeiten, sich an die Eigenschaften der zu vergleichenden Fahrzeuge zu erinnern. In dieser Studie war den Probanden die Referenz jedoch stets präsent, da sie diese immer zu Beginn des Manövers sowie auf den Transfers erleben konnten.

Die häufige Wiederholung der Referenz könnte dazu führen, dass sich die Probanden an diese Einstellung gewöhnen. Entsprechend der confirmation-expectation-theory (Oliver, 1980) findet ein kognitiver Abgleich zwischen der Erwartungshaltung und dem tatsächlichen Fahrzeugverhalten statt. Wird die Erwartung bestätigt, kann dies zu einer besseren Bewertung führen. Jedoch sind keine Auffälligkeiten bezüglich einer Bevorzugung der Referenz im Kontext der angrenzenden Varianten erkennbar. Dennoch wird für nachfolgende Versuche empfohlen, verschiedene Varianten als Referenz zu nutzen und diese beispielsweise zwischen den unterschiedlichen Sets zu variieren.

Weiterhin könnte das Versuchsdesign einen Einfluss auf die Bewertungen gehabt haben. Möglicherweise wurden die Unterschiede zwischen den Varianten durch den unmittelbaren Vergleich viel deutlicher wahrgenommen als dies im normalen Fahrbetrieb der Fall wäre. Den Probanden wurde mitgeteilt, dass ihnen sechs verschiedene Varianten präsentiert werden. Es wäre denkbar, dass sie aufgrund dieser Instruktion auch dann Unterschiede äußerten, wenn sie eigentlich keine wahrnahmen. Deshalb könnten die Differenzen zwischen den Varianten im Sinne der sozialen Erwünschtheit überschätzt sein (Crowne & Marlowe, 1960; Turner & Martin, 1984). Dem widerspricht jedoch, dass kaum Unterschiede zwischen der ersten und der wiederholten Bewertung gefunden wurden.

Versuchsfahrzeug

Im Gegensatz zu den vorherigen Versuchen wurde in diesem ein einziges Fahrzeug genutzt, sodass eine Beeinflussung durch unterschiedliche Hersteller ausgeschlossen werden konnte. Hinsichtlich der Übertragbarkeit

der Ergebnisse sind jedoch zwei Aspekte zu kritisieren. Zum einen wurde der Auflieger nicht beladen, da längere Beschleunigungsphasen mit den zur Verfügung stehenden Flächen nicht vereinbar gewesen wären. Zweitens wären die Manöver mit beladenem Auflieger schwerer zu beherrschen und somit sicherheitskritisch gewesen. Dies gilt insbesondere bei reduzierter Reibung, beispielsweise auf regennasser Fahrbahn.

Andererseits ist bei der Verwendung lediglich eines Fahrzeugs festzustellen, dass die Ergebnisse nur bedingt auf die Gesamtheit aller Fahrzeuge eines Segments übertragbar sind. Schließlich unterscheiden sich verschiedene Fahrzeuge hinsichtlich der Lenkübersetzung sowie der Fahrdynamik insgesamt (Gier- und Wankverhalten, Querdynamik). Da diese Faktoren das Lenkgefühl im weiteren Sinn beeinflussen (Wolf, 2009), muss vor einer Übertragung der Gestaltungsempfehlungen auf andere Fahrzeugtypen (bspw. Gliederzüge) eine erneute Validierung erfolgen.

Objektive Messdaten & Auswertung

Da der Fokus des vorliegenden Fahrerversuchs auf den Fahrer-Bewertungen lag, wurden die aufgezeichneten objektiven Messdaten nicht in ihrer Gänze ausgewertet. Beispielsweise wurden keine fahrerindividuellen Kennwerte der eingestellten Lenkwinkel und Lenkmomente berechnet und zu den Analysen herangezogen (Decker, 2009). Mit den aufgezeichneten Daten könnten beispielsweise auch multiple Regressionen mit weiteren objektiven Kennwerten durchgeführt werden (Riedel & Arbinger, 1997; Zschocke & Albers, 2008). Auch eine Analyse mit neuronalen Netzen wäre möglich (King, Crolla, Ash & Whitehead, 2002; Rothhämel, 2013).

Neben den fahrdynamischen Parameter haben noch weitere Fahrzeugparameter einen Einfluss auf das Lenkgefühl (Barthenheier, 2004). Dies gilt allgemein für die Sitzposition bzw. die Position des Fahrers zum Lenkrad (Wolf, 2009). Andererseits sind auch die Stellung des Lenkrads und die Position der Hände am Lenkrad entscheidend. Im Rahmen des Versuchs wurde keiner dieser Kennwerte erfasst, sodass die Bewertungen nicht in Bezug gesetzt werden können. Es kann somit nicht ermittelt werden, welchen Einfluss diese Parameter auf das Lenkgefühl hatten. Da jedoch keiner der Fahrer während der Fahrt eine Verstellung von Sitz oder Lenkrad vornahm, ist von einem kontinuierlichen Einfluss auszugehen.

Gestaltung der Varianten

Im Rahmen der Auswertung zu Hypothese 4 zeigte sich, dass die Probanden die Unterschiede zwischen allen Varianten deutlich wahrnahmen. Dies war gewünscht, um das Versuchskonzept zu validieren. Um noch detailliertere Gestaltungsempfehlungen ableiten zu können, hätten die dargestellten Unterschiede noch kleiner sein müssen. Dies hätte bei einer steigenden Anzahl an Varianten wiederum zu einer nicht unerheblichen Verlängerung des Versuchs geführt. Somit stellt die vorliegende Lösung einen Kompromiss aus detaillierten Ergebnissen und Versuchsökonomie dar.

Weiterhin war aufgrund der Vorgehensweise bei der technischen Realisierung der Varianten nur ein bestimmter Bereich der Lenkunterstützung abbildbar. Dass dieser Bereich jedoch nicht ganz ausreichend war, zeigte sich insbesondere beim Rangieren. Da hier die leichteste Variante bevorzugt wurde, kann nicht sichergestellt werden, dass eine noch leichtgängigere Variante nicht zu einer noch besseren Bewertung geführt hätte. Die Forderung

nach einem größeren untersuchten Bereich steht jedoch im Widerspruch zum vorher diskutierten Bedarf an einer feineren Auflösung. An dieser Stelle ergibt sich somit erneuter Forschungsbedarf für weitere Probandenstudien.

Ganz allgemein ist anzumerken, dass in dieser Studie lediglich die Variation der Lenkunterstützung untersucht wurde. Im Gegensatz zu anderen Studien (Ackert, 2008; Decker, 2009) wurde beispielsweise hinsichtlich der Lenkübersetzung keine systematische Variation vorgenommen. Somit wurde im Rahmen des Versuchs lediglich ein Aspekt des Lenkgefühls erfasst. Die Einbeziehung weiterer Faktoren hätte jedoch, wie oben angeführt, zu einer deutlichen Überschreitung des zeitlichen Rahmens geführt.

Übertragbarkeit der Ergebnisse auf reale Fahrsituationen

Um den unmittelbaren Vergleich der Varianten zu ermöglichen, wurden synthetische Manöver generiert. Nur so war ein direkter Vergleich mit der Referenz realisierbar. Insgesamt erwies sich dieses Vorgehen als zielführend, da die Probanden fundierte, differenzierte Abgleiche vornehmen konnten.

Es war den Fahrern andererseits aber nicht möglich, sich über einen längeren Zeitraum an eine Variante zu gewöhnen. Beispielsweise bei der Geradeausfahrt wurden die Varianten weniger als 20 Sekunden erprobt. Dieser Zeitraum erscheint sehr kurz und steht im Widerspruch zu den Anforderungen des Versuchskonzepts. Dennoch warteten die Probanden in den seltensten Fällen das Ende des Manövers ab, bis sie sich zu ihren Wahrnehmungen bzw. der Bewertung äußerten. Dies deckt sich mit Erkenntnissen von Barthenheier (2004), wonach Probanden bereits einige Sekunden nach dem Umschalten ein Urteil fällen und sich dieses über die Dauer des Manövers nicht verändert. Somit schien der angebotene Bewertungszeitraum ausreichend gewesen zu sein. Dessen ungeachtet kann eine etwas längere Erprobungsdauer besonders bei sehr leichtgängigen oder schwergängigen Varianten sinnvoll sein, da sich die Probanden erst an die unbekannte Situation gewöhnen müssen (Anand, 2014). Dies geht damit einher, dass Fahrer für normale Fahrsituationen die Abhängigkeit zwischen Lenkradwinkel und Fahrzeugreaktion verinnerlicht haben. Grobe Abweichungen von erlernten Zusammenhängen führen jedoch zu Schwierigkeiten, beispielsweise bei der Kurvenfahrt (Schmidtke, 1993). Somit könnte eine längere Erprobung zu einer Habituation führen und dadurch voreilige negative Bewertungen extremer Varianten verhindern. Für zukünftige Bewertungen könnte angedacht werden, die Instruktion so anzupassen, dass Probanden den gesamten Zeitraum zur Bewertung nutzen müssen.

Abschließend könnte die Übertragbarkeit der Ergebnisse davon abhängen, inwiefern die Fahrer die Manöver auf den täglichen Fahrbetrieb abstrahierten. Während die Geradeausfahrt und die Rangiersituation sehr realitätsnah waren, war beispielsweise der Slalom ein eher synthetisches Manöver. Somit ist also nicht sichergestellt, dass die Probanden ihre Wahrnehmungen auf reale Situationen übertragen konnten. Dies könnte dadurch unterstützt worden sein, dass die gewählten Geschwindigkeiten niedriger waren als sie im realen Straßenverkehr vorzufinden wären. Für den Slalom und den doppelten Spurwechsel, welche landstraßenähnliche Fahrmanöver darstellen, wären 60 km/h entsprechend der zulässigen Höchstgeschwindigkeit auf deutschen Landstraßen für Lkw zu wählen gewesen. Für die Geradeausfahrt, welche einer Fahrsituation auf der Autobahn entspricht, wären 80 km/h als Höchstgeschwindigkeit angezeigt.

5. Beantwortung der Forschungsfragen

Im Rahmen dieser Arbeit wurde ein Versuchskonzept entwickelt, mit welchem das Lenkgefühl im schweren Nutzfahrzeug von Normalfahrern bewertet werden kann. Dieses wurde in zwei Fahrversuchen evaluiert und weiter optimiert. Anschließend diente das Versuchskonzept dazu, verschiedene Parametrierungen der Lenkunterstützung hinsichtlich des resultierenden Lenkgefühls zu bewerten und zu vergleichen. Auf Basis dieser Arbeiten und der daraus gewonnenen Erkenntnisse werden nachfolgend die vier Forschungsfragen beantwortet (Tabelle 45).

Tabelle 45. Forschungsfragen, dazugehörige Forschungsarbeiten und Ergebnisse.

	Forschungsfrage	Forschungsarbeit	Ergebnis
1	Mit welchem Versuchskonzept können Normalfahrer das Lenkgefühl am schweren Nutzfahrzeug bewerten?	<ul style="list-style-type: none">• Entwicklung eines Versuchskonzepts (Kapitel 3.1)• Evaluation in Studie I (Kapitel 3.2)• Evaluation in Studie II (Kapitel 3.3)	Versuchskonzept (Kapitel 3.4)
2	Kann das Lenkgefühl verschiedener Fahrzeuge unter Anwendung eines geeigneten Versuchskonzepts in realitätsnahen Fahrversuchen evaluiert werden?	<ul style="list-style-type: none">• Studie I (Kapitel 3.2)• Studie II (Kapitel 3.3)	Nicht bestätigt
3	Können unter Anwendung eines geeigneten Versuchskonzepts Erkenntnisse zur Objektivierung des Lenkgefühls aus Sicht der Normalfahrer generiert werden?	<ul style="list-style-type: none">• Studie III (Kapitel 4)	Bestätigt
4	Wie können Normalfahrer das Lenkgefühl am schweren Nutzfahrzeug zuverlässig bewerten?	<ul style="list-style-type: none">• Studie I (Kapitel 3.2)• Studie II (Kapitel 3.3)• Studie III (Kapitel 4)	Unter Anwendung des Versuchskonzepts (Kapitel 3.4)

Forschungsfrage 1: Mit welchem Versuchskonzept können Normalfahrer das Lenkgefühl am schweren Nutzfahrzeug bewerten?

Auf Basis von Interviewstudien, einem explorativen Fahrversuch und Experten-Interviews wurde ein Fragebogen für Normalfahrer entwickelt und in ein entsprechendes Versuchskonzept übertragen. Die Anforderungen an den Fragebogen wurden aus der Literatur abgeleitet. Der Schwerpunkt lag auf einer hohen Verständlichkeit für Normalfahrer bei gleichzeitiger Detailliertheit der Bewertungskriterien. Deshalb wurden diese zusätzlich mit erklärenden Beschreibungen versehen. Ausgehend von den Fahrversuchen und umfassenden Literaturbelegen wurde eine vollständig mit verbalen Benennungen versehene 7-stufige Likert-Skala zur Bewertung ausgewählt. Der so entstandene Fragebogen wurde in einem iterativen Prozess in ein Versuchskonzept übertragen. Allem voran ist hier die Zuordnung relevanter Fahrsituationen zu den einzelnen Bewertungskriterien zu nennen.

Außerdem wurden Anforderungen zu Versuchsdesign, Operationalisierung, Bewertungsvorgehen und Stichprobe zusammengetragen. Das finale Versuchskonzept ist in Kapitel 3.4 zu finden.

Forschungsfrage 2: Kann das Lenkgefühl verschiedener Fahrzeuge unter Anwendung eines geeigneten Versuchskonzepts in realitätsnahen Fahrversuchen evaluiert werden?

In den Studien I und II wurden verschiedene Fahrzeuge im Straßenverkehr hinsichtlich ihres Lenkgefühls absolut bewertet. In Studie I wurden zwei Sattelzüge unmittelbar hintereinander evaluiert. Die vermuteten Unterschiede im Lenkgefühl wurden nur zum Teil aufgedeckt. Die Ergebnisse legten eine Beeinflussung der Bewertungen durch das üblicherweise von den Probanden gefahrene Fahrzeug nahe. In Studie II wurden vier verschiedene Sattelzüge im Abstand mehrerer Tage erprobt. Auch hier wurden die vermuteten Unterschiede nur teilweise aufgedeckt. Dies könnte im zeitlichen Abstand zwischen den einzelnen Fahrsituationen bzw. Versuchsfahrten begründet sein. Alles in allem ist die Frage, ob das Lenkgefühl von Normalfahrern mithilfe des entwickelten Versuchskonzepts in realitätsnahen Fahrversuchen evaluiert werden kann, zu verneinen. Schließlich gab es in keinem der beiden Versuche belastbare Ergebnisse, wonach Normalfahrer Unterschiede im Lenkgefühl zwischen verschiedenen Lkw während der Fahrt im Straßenverkehr wahrnehmen.

Forschungsfrage 3: Können unter Anwendung eines geeigneten Versuchskonzepts Erkenntnisse zur Objektivierung des Lenkgefühls aus Sicht der Normalfahrer generiert werden?

In Studie III wurden verschiedene Varianten der Lenkungsauslegung auf der Teststrecke evaluiert und relativ mit einer Referenz verglichen. Die Analyse der Niveau-Bewertungen zeigte deutlich, dass die Fahrer die objektiv vorhandenen Unterschiede wahrnahmen. Auch die Gefallen-Bewertungen der Varianten fielen unterschiedlich aus, was sich in den entsprechenden Gestaltungsempfehlungen widerspiegelte. Ferner konnten die Niveau- und Gefallen-Bewertungen aus den zugrunde liegenden objektiven Parametern des Lenksystems vorhergesagt werden. Dies galt jedoch bei den Gefallen-Bewertungen für die einzelnen Parameter in unterschiedlichem Maße. Für einige Bewertungskriterien wurden sehr enge Zusammenhänge zwischen den objektiven Kennwerten und den Gefallen-Bewertungen gefunden. Für andere wiederum waren die Zusammenhänge schwächer, was für individuelle Vorlieben bezüglich dieser Kriterien spricht. Somit kann die Forschungsfrage, ob mithilfe des entwickelten Versuchskonzepts Erkenntnisse zur Objektivierung des Lenkgefühls aus Sicht der Normalfahrer generiert werden können, bejaht werden.

Forschungsfrage 4: Wie können Normalfahrer das Lenkgefühl am schweren Nutzfahrzeug zuverlässig bewerten?

Zur Beantwortung dieser Forschungsfrage wurden die Ergebnisse aller drei Studien herangezogen. Diese zeigten, dass die Brauchbarkeit und Zuverlässigkeit der subjektiven Bewertungen der Fahrer von der Versuchsumgebung bzw. der Variation der objektiven Parameter abhing. Bei der Evaluation verschiedener Fahrzeuge im Straßenverkehr konnten keine belastbaren Daten gesammelt werden. Außerdem waren die Urteile zum Teil nicht stabil gegenüber Wiederholungen. Möglicherweise wurden die Bewertungen also durch andere Aspekte als das Lenkgefühl beeinflusst. Andererseits könnte auch der zeitliche Abstand zwischen den Bewertungen die Vergleichbarkeit beeinträchtigt haben. Diese beiden Aspekte wurden in der letzten Studie kontrolliert und die Zuverlässigkeit erneut systematisch erfasst. Bei der blinden Wiederholung einer Variante innerhalb des

Fahrmanövers waren bis auf wenige Ausnahmen keine Unterschiede in den Bewertungen zu finden. Ferner konnten auch die vermuteten Zusammenhänge zwischen den variierten objektiven Parametern und den subjektiven Bewertungen des Lenkgefühls aufgedeckt werden. Aus den drei Studien geht somit hervor, dass Normalfahrer zuverlässig bewerten können, wenn verschiedene Varianten des Lenkgefühls in einem Fahrzeug auf der Teststrecke evaluiert werden. Die Bewertung fand dabei relativ und in Bezug auf eine Referenz statt. Diese Randbedingungen entsprechen den Anforderungen zur Operationalisierung im finalen Versuchskonzept.

6. Ausblick

In der vorliegenden Arbeit wurde ein Versuchskonzept entwickelt, um das Lenkgefühl am schweren Nutzfahrzeug für Normalfahrer bewertbar zu machen. Dieser Ansatz wurde anschließend dazu genutzt, Erkenntnisse zum Lenkgefühl zu generieren. Im Folgenden wird ein Ausblick für weitere Forschungsarbeiten gegeben.

Zuerst soll auf das Versuchskonzept eingegangen werden. Hier ist einerseits anzumerken, dass hinsichtlich des Fragebogens noch Forschungsbedarf nötig ist. Obwohl der Fragebogen bereits in drei Studien verwendet wurde, konnte er nie an einer größeren Stichprobe validiert werden. Die durchgeführten Analysen sollten deshalb mit einer größeren Datengrundlage wiederholt werden. Es wird jedoch unbedingt empfohlen, die Daten im Rahmen einer Erprobung im Fahrzeug zu erheben und bspw. nicht auf Interviews oder online-Befragungen zurückzugreifen. Hieraus ergibt sich die Problematik, dass für eine große Datenbasis ein entsprechend großer Zeitaufwand nötig ist.

Neben der weiteren Optimierung ist auch eine Ausweitung des Versuchskonzepts auf andere Fragestellungen der subjektiven Bewertung denkbar. Die Methode selbst ist mit Ausnahme des Fragebogens sicherlich mit geringen Abänderungen auch auf die Bewertung anderer Fahrzeuge übertragbar. Mit kleineren Anpassungen kann der Fragebogen auch für andere schwere Nutzfahrzeuge als den Lkw genutzt werden, z. B. für Busse. Zur Anwendung auf den Pkw wird der Fragebogen jedoch nicht empfohlen, da insbesondere hinsichtlich der Kategorien Komfort und Sportlichkeit andere Anforderungen bestehen. Hier ergibt sich jedoch zusätzlich die Notwendigkeit, das Fahrgefühl der Passagiere adäquat zu beachten. Schließlich sollte hier das komfortable Lenkgefühl des Fahrers zu für die Passagiere angenehmen Fahrzeugreaktionen führen. Ähnliche Fragestellungen treten auch bei der Entwicklung hin zum hochautomatisierten Fahren auf. Wie beim Bus sollte auch hier ein für die Passagiere angenehmes Lenkverhalten realisiert werden.

Es ergab sich jedoch auch aus den Ergebnissen weiterer Forschungsbedarf, beispielsweise in Bezug auf die dargestellten Varianten. Da die meisten Probanden die Unterschiede zwischen ihnen wahrnahmen, könnten diese in einem zukünftigen Versuch kleiner ausfallen, um noch detailliertere Bewertungen zu erzielen. Andererseits zeigte sich, dass mit den Varianten ein noch breiterer Bereich des vom Fahrer aufzuwendenden Lenkradmoments hätte dargestellt werden müssen. Da diese Forderungen jedoch zu einer noch längeren Versuchsdauer führen würden, muss ein alternatives Konzept entwickelt werden. Entgegen der Bewertung verschiedener vorher entwickelter Varianten könnte das Versuchskonzept so angepasst werden, dass die Probanden selbst Veränderungen in den Einstellungen vornehmen können.

Unabhängig von diesen Anpassungen basiert das entwickelte Versuchskonzept darauf, verschiedene Varianten des Lenkgefühls möglichst unmittelbar und kompakt zu bewerten und zu vergleichen. Deshalb werden die einzelnen Varianten lediglich für einen sehr kurzen Zeitraum erprobt. Deshalb ist nicht sichergestellt, dass die Probanden ihre Bewertung auf den Fahralltag abstrahieren konnten. Die so abgeleiteten Gestaltungsempfehlungen sollten jedoch nicht ohne weitere Überprüfung zur Serienreife gebracht werden. Aus diesen Gründen sollten auf Basis der Empfehlungen zwei oder drei Varianten in einem Fahrzeug umgesetzt werden. Diese sollten anschließend im normalen Fahrbetrieb evaluiert werden, beispielsweise für die Dauer eines Arbeitstages.

Schlussendlich verdeutlicht diese Forschungsarbeit, dass Normalfahrer unter Nutzung des entwickelten Versuchskonzepts als zusätzliche Quelle im Auslegungsprozess genutzt werden können. Für nachfolgende Studien wird deshalb empfohlen, die verwendete Methode auch für die Untersuchung weiterer Parameter zu nutzen. Mit Steer-by-Wire-Systemen wäre es mit vergleichsweise geringem Aufwand möglich, mehrere Parameter gleichzeitig und in einem breiten Spektrum zu variieren. Neben dem Lenkmoment könnten so auch Lenkübersetzung, Dämpfung, Reibung etc. systematisch verändert werden. Erst wenn alle Parameter gemeinsam systematisch variiert und evaluiert werden, liegt ein ganzheitliches Bild der Zusammensetzung des Lenkgefühls vor. Die entwickelte Methode bietet hierbei die Grundlage für die Erfassung der subjektiven Bewertungen.

7. Literaturverzeichnis

- Aaker, J. L. (1997). Dimensions of Brand Personality. *Journal of Marketing Research*, 34, 347-356.
- Abendroth, B. & Bruder, R. (2015). Die Leistungsfähigkeit des Menschen für die Fahrzeugführung. In H. Winner, S. Hakuli, F. Lotz & C. Singer (Hrsg.), *Handbuch Fahrerassistenzsysteme* (S. 3-15). Wiesbaden: Springer Fachmedien.
- Ackert, H. (2008). *Ergonomische Bewertung der Sicherheit beim Kraftfahrzeuglenken: Untersuchungen zum Zusammenhang von subjektiver Bewertung und realisierten Fahrleistungen bei einer parameterabhängigen Lenkungsgestaltung*. Dissertation, Technische Universität Darmstadt. Darmstadt.
- Alwin, D. F. & Krosnick, J. A. (1991). The Reliability of Survey Attitude Measurement. The Influence of Question and Respondent Attributes. *Sociological Methods & Research*, 20 (1), 139-181.
- Anand, S. (2014). *Personalizing steering experience using steer-by-wire systems*. Dissertation, Technische Universiteit Eindhoven. Eindhoven.
- Anand, S., Terken, J. & Hogema, J. (2011). Individual differences in preferred steering effort for steer-by-wire systems. In M. Tscheligi (Hrsg.), *Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (S. 55-62). New York, NY: ACM.
- Badiru, I. A. (2014). Customer Focus in EPS Steering Feel Development. *SAE International Journal of Passenger Cars - Mechanical Systems*, 7 (3), 1009-1015.
- Baier, J. (2012). *Trends im Straßengüterverkehr - Aktueller Status und Meinungen der Berufskraftfahrer in Deutschland*. Furtwangen: Hochschule Furtwangen University. Verfügbar unter http://www.hs-furtwangen.de/fileadmin/user_upload/Marketing_PR/Dokumente/Webredaktion/Studie_Berufskraftfahrer_2011.pdf
- Barthenheier, T. (2004). *Potenzial einer fahrertyp- und fahrsituationsabhängigen Lenkradmomentgestaltung*. Dissertation, Technische Universität Darmstadt. Darmstadt.
- Barthenheier, T. (2006). Potenzial einer individuellen Lenkmomentgestaltung. In K. Becker (Hrsg.), *Subjektive Fahreindrücke sichtbar machen III. Korrelation zwischen objektiver Messung und subjektiver Beurteilung von Versuchsfahrzeugen und -komponenten* (Bd. 56, S. 50-73). Renningen: expert-Verl.
- Barthenheier, T. & Winner, H. (2003). *Das persönliche Lenkgefühl*. Vortrag auf der Fahrwerk.tech 2003, München.
- Bass, B. M., Cascio, W. F. & O'Connor, E. J. (1974). Magnitude estimations of expressions of frequency and amount. *Journal of Applied Psychology*, 59 (3), 313-320.
- Becker, R. (2011). Vorsicht: Suchtgefahr. Fahrer testen: Scania R 730 Topline. *Fernfahrer* (1), 24-31.
- Beggan, J. K. (1992). On the social nature of nonsocial perception. The mere ownership effect. *Journal of Personality and Social Psychology*, 62 (2), 229-237.
- Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57 (1), 289-300.
- Bertollini, G. P. & Hogan, R. M. (1999). Applying Driving Simulation to Quantify Steering Effort Preference as a Function of Vehicle Speed. In *International Congress & Exposition* (SAE Technical Paper Series). SAE International.
- Bianchi, I., Savardi, U. & Burro, R. (2011). Perceptual ratings of opposite spatial properties: Do they lie on the same dimension? *Acta Psychologica*, 138 (3), 405-418.

- Black, J. (2010). *Vehicle Steering Systems - Hardware-in-the-Loop Simulator, Driving Preferences, and Vehicle Intervention*. Dissertation, Clemson University.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Mit 87 Tabellen* (Springer-Lehrbuch : Bachelor, Master, 4., überarb. Aufl.). Heidelberg: Springer.
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler. Mit 163 Tabellen* (Springer-Lehrbuch, 7., vollst. überarb. u. erw. Aufl.). Berlin u.a.: Springer.
- Braess, H.-H. (2001). Steering system, steering system properties, vehicle dynamics. The progress of the last 50 years, an outlook into the future. In *Reifen, Fahrwerk, Fahrbahn. Tagung Hannover, 18. und 19. Oktober 2001* (VDI-Berichte, 1632, Buch, S. 13-55). Düsseldorf: VDI-Verl.
- Braess, H.-H. (2004). Die schwierige Übung des richtigen Kurses. *Frankfurter Allgemeine Zeitung*.
- Braess, H.-H. & Seiffert, U. (Hrsg.). (2013). *Vieweg Handbuch Kraftfahrzeugtechnik*. Wiesbaden: Springer Fachmedien.
- Breuer, J. (2000). Beurteilung des Fahrverhaltens im Fahrsimulator der DaimlerChrysler AG und in realen Fahrzeugen. In K. Becker (Hrsg.), *Subjektive Fahreindrücke sichtbar machen. Korrelation zwischen CAE-Berechnung, Versuch und Messung von Versuchsfahrzeugen und -komponenten* (S. 43-53). Renningen-Malmsheim: expert-Verl.
- Bruder, R., Abendroth, B. & Landau, K. (2007). Zum Nutzen von Fahrversuchen für die Gestaltung. In R. Bruder (Hrsg.), *Wie objektiv sind Fahrversuche? Darmstädter Kolloquium Mensch & Fahrzeug, Technische Universität Darmstadt, 14./15. März 2007* (S. 79-95). Stuttgart: Ergonomia-Verl.
- Bubb, H. (2002). Der Fahrprozess - Informationsverarbeitung durch den Fahrer. In VDA (Hrsg.), *4. Technischer Kongress 2002. Sicherheit durch Elektronik / Fahrzeugsicherheit, Elektronik* (S. 19-31). VDA.
- Bubb, H. (2003). Wie viele Probanden braucht man für allgemeine Erkenntnisse aus Fahrversuchen? In K. Landau & H. Winner (Hrsg.), *Fahrversuche mit Probanden - Nutzwert und Risiko. Darmstädter Kolloquium Mensch & Fahrzeug, Technische Universität Darmstadt, 3./4. April 2003* (Fortschritt-Berichte VDI Reihe 12, Verkehrstechnik/Fahrzeugtechnik, Bd. 557, Als Ms. gedr, S. 26-39). Düsseldorf: VDI-Verl.
- Bubb, H. & Schmidtke, H. (1993). Systemergonomie. In H. Schmidtke (Hrsg.), *Ergonomie* (3., neubearb. und erw. Aufl., S. 305-458). München: Hanser.
- Buchner, C. (2008). *Die Wirkung von Produktvielfalt auf die Markenstärke. Theoretische Überlegungen, empirische Befunde und Handlungsempfehlungen am Beispiel des Premiumsegments des Automobilmarktes* (Schriftenreihe Schwerpunkt Marketing, Bd. 71). Zugl.: München, Univ., Diss., 2007. München: FGM-Verl.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (PS Psychologie, 3., aktualisierte und erw. Aufl.). München u.a.: Pearson Studium.
- Buschardt, B. (2003). *Synthetische Lenkmomente* (ZMMS-Spektrum, Bd. 16, Als Ms. gedr). Techn. Univ., Diss.-Berlin. Düsseldorf: VDI-Verl.
- Bush, L. K., Hess, U. & Wolford, G. (1993). Transformations for within-subject designs: a Monte Carlo investigation. *Psychological Bulletin*, 113 (3), 566-579.
- Caird, J. K. & Horrey, W. J. (2011). Twelve practical and useful questions about driving simulation. In D. L. Fisher (Ed.), *Handbook of driving simulation for engineering, medicine, and psychology* (p. Kap. 5). Boca Raton: CRC Press.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56 (2), 81-105.

- Chen, D. (1997). *Subjective And Objective Vehicle Handling Behaviour*, University of Leeds. Leeds.
- Chen, D. C. & Crolla, D. A. (1998). Subjective and objective measures of vehicle handling. *Vehicle System Dynamics*, 29 (sup1), 576-597.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed.). Hillsdale NJ u.a.: Erlbaum.
- Cook, R. D. & Weisberg, S. (1982). *Residuals and influence in regression* (Monographs on statistics and applied probability, [18], 1. publ). New York NY u.a.: Chapman and Hall.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78 (1), 98-104.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16 (3), 297-334.
- Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24 (4), 349-354.
- Cummins, C. (2012). *Critical Values for the Durbin-Watson Test*, Stanford University. Verfügbar unter <http://web.stanford.edu/~clint/bench/dwcrit.htm>
- Dang, J., Chen, H., Gao, B., Li, Q., Li, M., Watanabe, T. et al. (2014). Optimal Design of On-Center Steering Force Characteristic Based on Correlations between Subjective and Objective Evaluations. *SAE International Journal of Passenger Cars - Mechanical Systems*, 7 (3), 992-1001.
- Dang, J., Chen, H., Li, Q., Watanabe, T., Hayama, R. & Lou, L. (2015). A method to set the target on-centre steering force characteristic. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 230 (12), 1670-1686.
- Decker, M. (2009). *Zur Beurteilung der Querdynamik von Personenkraftwagen*. Dissertation, Technische Universität München. München.
- Dettki, F. (2005). *Methoden zur objektiven Bewertung des Geradeauslaufs von Personenkraftwagen*. Dissertation, Universität Stuttgart. Stuttgart.
- Dijkstra, W. (1987). Interviewing Style and Respondent Behavior. An Experimental Study of the Survey-Interview. *Sociological Methods & Research*, 16 (2), 309-334.
- Durbin, J. & Watson, G. S. (1951). Testing for Serial Correlation in Least Squares Regression: II. *Biometrika*, 38 (1/2), 159.
- Dziuban, C. D. & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? *Psychological Bulletin*, 81 (6), 358-361.
- Ellmann, A. (2003). *Vielfahrer als Zielgruppe für Verkehrssicherheitsaktionen*. Presseseminar "Vielfahrer in Deutschland", Münster.
- Engelbrecht, A. (2013). *Fahrkomfort und Fahrspaß bei Einsatz von Fahrerassistenzsystemen*. Dissertation, Humboldt Universität, Berlin. Norderstedt: Grin Verlag GmbH.
- ISO, 9241-210:2010 (2011). *Ergonomie der Mensch-System-Interaktion - Teil 210: Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme*.
- Farrer, D. G. (1993). An Objective Measurement Technique for the Quantification of On-Centre Handling Quality. In *International Congress & Exposition* (SAE Technical Paper Series). SAE International.
- Fastenmeier, W. (2005). *Das Fahrverhalten in Kreuzungen: Welche Fehler begehen ältere Autofahrer?* Deutscher Psychologentag, Potsdam.

- Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, 41 (4), 1149-1160.
- Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3. A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39 (2), 175-191.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Los Angeles: Sage Publ.
- Friborg, O., Martinussen, M. & Rosenvinge, J. H. (2006). Likert-based vs. semantic differential-based scorings of positive psychological constructs. A psychometric comparison of two versions of a scale measuring resilience. *Personality and Individual Differences*, 40 (5), 873-884.
- Gies, S. & Marusic, Z. (2000). Das Lenkgefühl - Merkmale der subjektiven und objektiven Beschreibung. In K. Becker (Hrsg.), *Subjektive Fahrindrücke sichtbar machen. Korrelation zwischen CAE-Berechnung, Versuch und Messung von Versuchsfahrzeugen und -komponenten* (S. 67-84). Renningen-Malmsheim: expert-Verl.
- Gil Gómez, G. L., Nybacka, M., Bakker, E. & Drugge, L. (2015). Findings from subjective evaluations and driver ratings of vehicle dynamics. Steering and handling. *Vehicle System Dynamics*, 53 (10), 1416-1438.
- Google Maps. (2014). *Versuchsstrecke Studie I*. Zugriff am 21.12.2014. Verfügbar unter <https://www.google.de/maps/@48.6560681,8.9823076,13z?hl=de>
- Google Maps. (2016). *Versuchsstrecke Studie II*. Zugriff am 23.02.2016. Verfügbar unter <https://www.google.de/maps/@48.7157356,9.1886446,12z?hl=de>
- Greenhouse, S. W. & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24 (2), 95-112.
- Greenleaf, E. A. (1992). Improving Rating Scale Measures by Detecting and Correcting Bias Components in Some Response Styles. *Journal of Marketing Research*, 29 (2), 176-188.
- Grünig, G. (2013). *Vergleichstest: Ein Quartett mit Sechsappeal*, Trucker.de. Zugriff am 28.09.2016. Verfügbar unter <http://www.trucker.de/vergleichstest-ein-quartett-mit-sechsappeal-1251743.html>
- Guadagnoli, E. & Velicer, W. F. (1988). Relation to sample size to the stability of component patterns. *Psychological Bulletin*, 103 (2), 265-275.
- Gutjahr, D. (2013). *Objektive Bewertung querdynamischer Reifeneigenschaften im Gesamtfahrzeugversuch*. Dissertation, Karlsruher Institut für Technologie. Karlsruhe.
- Harrer, M. (2007). *Characterisation of Steering Feel*. Dissertation, University of Bath.
- Harrer, M. & Pfeffer, P. (2013). Lenkgefühl, Interaktion. In P. Pfeffer & M. Harrer (Hrsg.), *Lenkungshandbuch* (S. 125-144). Wiesbaden: Springer Fachmedien.
- Harrer, M., Stickel, T. & Pfeffer, P. (2005). Automatisierung fahrdynamischer Messungen. In *Reifen, Fahrwerk, Fahrbahn. Markt- und kundengerechte Innovation ; [Tagung Hannover, 25. und 26. Oktober 2005 ; 10. Fachtagung "Reifen, Fahrwerk, Fahrbahn"]* (VDI-Berichte, 1912, CD-ROM, S. 429-449). Düsseldorf: VDI-Verl.
- Heißing, B. & Brandl, H. J. (2002). *Subjektive Beurteilung des Fahrverhaltens* (Vogel-Fachbuch, 1. Aufl.). Würzburg: Vogel.
- Heißing, B. & Ersoy, M. (2011). *Chassis Handbook. Fundamentals, Driving Dynamics, Components, Mechatronics, Perspectives* (1. Aufl.). s.l.: Vieweg+Teubner (GWV).
- Heißing, B., Ersoy, M. & Gies, S. (Hrsg.). (2013). *Fahrwerkhandbuch*. Wiesbaden: Springer Fachmedien.
- Hempel, T. (2001). *Untersuchungen zur Korrelation auditiver und instrumenteller Messergebnisse für die Bewertung von Fahrzeuginnenraumgeräuschen als Grundlage eines Beitrags zur Klassifikation von*

- Hörereignissen* (Mensch-Maschine-Kommunikation). Zugl.: Berlin, Techn. Univ., Diss., 2001. München: Utz Wiss.
- Henrichfreise, H., Jusseit, J. & Niessen, H. (2003). Optimale Regelung einer elektromechanischen Servolenkung. In *5. VDI Mechatronik Tagung. Innovative Produktentwicklung*.
- Hiraoka, T., Hioki, S. & Kumamoto, H. (2008). Fundamental research on driver's haptic characteristics concerning steering reactive torque. In *SICE 2008 - 47th Annual Conference of the Society of Instrument and Control Engineers of Japan* (S. 884-889).
- Hutcheson, G. & Sofroniou, N. (2009). *The multivariate social scientist. Introductory statistics using generalized linear models*. London: Sage Publications.
- Huynh, H. & Feldt, L. S. (1976). Estimation of the Box Correction for Degrees of Freedom from Sample Data in Randomized Block and Split-Plot Designs. *Journal of Educational and Behavioral Statistics*, 1 (1), 69-82.
- Jürgensohn, T. & Kolrep, H. (2006). Objektivierung der Wahrnehmung von Zitterbewegungen bei Cabrios. In K. Becker (Hrsg.), *Subjektive Fahrindrücke sichtbar machen III. Korrelation zwischen objektiver Messung und subjektiver Beurteilung von Versuchsfahrzeugen und -komponenten* (Bd. 56, S. 108-117). Renningen: expert-Verl.
- Jürgensohn, T. & Timpe, K.-P. (Hrsg.). (2001). *Kraftfahrzeugführung*. Berlin: Springer.
- Kallus, K. W. (2010). *Erstellung von Fragebogen*. Stuttgart: UTB.
- Karch, G. & Grüner, S. (2007). Mechatronische Lenksysteme (Mechatronic Steering Systems). *at – Automatisierungstechnik*, 55 (6).
- Kieruj, N. D. & Moors, G. (2010). Variations in Response Style Behavior by Response Scale Format in Attitude Research. *International Journal of Public Opinion Research*, 22 (3), 320-342.
- Kim, J. (2011). Objectification of on-center handling characteristics based on spring-mass-damper system. *International Journal of Automotive Technology*, 12 (6), 857-864.
- Kim, J. & Yoon, Y.-S. (2015). Objectification of steering feel around straight-line driving for vehicle/tyre design. *Vehicle System Dynamics*, 53 (2), 197-214.
- Kim, J.-H. & Song, J.-B. (2002). Control logic for an electric power steering system using assist motor. *Mechatronics*, 12 (3), 447-459.
- Kim, J.-O. & Mueller, C. W. (1978). *Factor analysis. Statistical methods and practical issues* (Sage university papers : Quantitative applications in the social sciences, Bd. 14, 26. [print.]. Newbury Park Calif. u.a.: Sage Publ.
- Kim, K. I. & van de Wiel, M. A. (2008). Effects of dependence in high-dimensional multiple testing problems. *BMC bioinformatics*, 9, 114.
- King, R. P., Crolla, D. A., Ash, H. A. S. & Whitehead, J. (2002). Identification of Subjective-Objective Vehicle Handling Links Using Neural Networks for the Foresight Vehicle. In *SAE 2002 World Congress & Exhibition* (SAE Technical Paper Series). SAE International.
- Kline, P. (2000). *The handbook of psychological testing* (2. ed.). London u.a.: Routledge.
- Knappe, G. (2009). *Empirische Untersuchungen zur Querregelung in Fahrsimulatoren*. Dissertation, Friedrich-Alexander-Universität. Erlangen-Nürnberg.
- Koch, T. (2010). *Untersuchungen zum Lenkgefühl von Steer-by-Wire Lenksystemen*. Dissertation, Technische Universität München. München.

- Koide, M. & Kawakami, S. (1988). Analysis of „Steering Feel“ Evaluation in Vehicles with Power Steering. *JSAE Review*, 36-42.
- Koschinsky, B. (2014). *Fahrdynamik Gesamtfahrzeug. Interner technischer Bericht*. Stuttgart: Daimler AG.
- Kraft, C. (2011). *Gezielte Variation und Analyse des Fahrverhaltens von Kraftfahrzeugen mittels elektrischer Linearaktuatoren im Fahrwerksbereich* (Karlsruher Schriftenreihe Fahrzeugsystemtechnik, Bd. 5). Karlsruhe: KIT Scientific Publishing.
- Kromrey, H. (1994). *Empirische Sozialforschung. Modelle und Methoden der Datenerhebung und Datenauswertung* (Uni-Taschenbücher, Bd. 1040, 6. revidierte Auflage). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Krüger, H.-P. & Neukum, A. (2001). Bewertung von Handlingeigenschaften - zur methodischen und inhaltlichen Kritik des korrelativen Forschungsansatzes. In T. Jürgensohn & K.-P. Timpe (Hrsg.), *Kraftfahrzeugführung* (S. 245-262). Berlin: Springer.
- Kudritzki, D. (2000). Möglichkeiten der Objektivierung subjektiver Beurteilungen des Fahrverhaltens. In K. Becker (Hrsg.), *Subjektive Fahreindrücke sichtbar machen. Korrelation zwischen CAE-Berechnung, Versuch und Messung von Versuchsfahrzeugen und -komponenten* (S. 11-26). Renningen-Malmsheim: expert-Verl.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (Grundlagen Psychologie, 6. Aufl., Studienausg.). Weinheim: Beltz.
- Likert, R. (1932). A technique for the measurement of attitudes. In R. S. Woodworth (Hrsg.), *Archives of Psychology* (Bd. 140, S. 5-55).
- Lombardi, C. M. & Hurlbert, S. H. (2009). Misprescription and misuse of one-tailed tests. *Austral Ecology*.
- Lozano, L. M., García-Cueto, E. & Muñiz, J. (2008). Effect of the Number of Response Categories on the Reliability and Validity of Rating Scales. *Methodology*, 4 (2), 73-79.
- MacCallum, R. C., Widaman, K. F., Zhang, S. & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4 (1), 84-99.
- Mandhata, U. B., Jensen, M. J., Wagner, J. R., Switzer, F. S., Dawson, D. M. & Summers, J. D. (2012). Evaluation of a customizable haptic feedback system for ground vehicle steer-by-wire interfaces. In *2012 American Control Conference - ACC 2012* (S. 2781-2787).
- Mansfield, N. & Whiting-Lewis, E. (2004). Low Frequency Lateral Acceleration and Subjective Ratings of Acceleration Intensity and Driving Confidence in Production Cars. *Low Frequency Noise, Vibration and Active Control*, 23 (4), 221-230.
- Mitschke, M. & Wallentowitz, H. (2014). Kreisfahrt bei konstanter Fahrgeschwindigkeit. In M. Mitschke & H. Wallentowitz (Hrsg.), *Dynamik der Kraftfahrzeuge* (S. 625-655). Wiesbaden: Springer Fachmedien.
- Moors, G. (2008). Exploring the effect of a middle response category on response style in attitude measurement. *Quality & quantity*, 42 (6), 779-794.
- Moosbrugger, H. & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion* (2. Aufl.). Berlin, Heidelberg: Springer.
- Müller, S. (2004). Zukünftige verbrauchsarme Servolenkungen für vollständige Steer-By-Wire-Funktionalität. *ATZ - Automobiltechnische Zeitschrift*, 106 (4), 310-318.
- Mundfrom, D. J., Shaw, D. G. & Ke, T. L. (2005). Minimum Sample Size Recommendations for Conducting Factor Analyses. *International Journal of Testing*, 5 (2), 159-168.

- Muñiz, J., García-Cueto, E. & Lozano, L. M. (2005). Item format and the psychometric properties of the Eysenck Personality Questionnaire. *Personality and Individual Differences*, 38 (1), 61-69.
- Neukum, A., Krüger, H.-P. & Schuller, J. (2001). Der Fahrer als Messinstrument für fahrdynamische Eigenschaften? In VDI-Gesellschaft Fahrzeug- und Verkehrstechnik (Hrsg.), *Der Fahrer im 21. Jahrhundert* (VDI-Berichte, Bd. 1613, S. 13-32). Düsseldorf: VDI-Verlag.
- Neukum, A., Paulig, J., Frömmig, L. & Henze, R. (2009). Untersuchung zur Wahrnehmung von Lenkmomenten bei Pkw. *FAT - Schriftenreihe* (222).
- Newberry, A. C., Griffin, M. J. & Dowson, M. (2007). Driver perception of steering feel. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 221 (4), 405-415.
- Nybacka, M., He, X., Gil Gómez, G. L., Bakker, E. & Drugge, L. (2014). Links between subjective assessments and objective metrics for steering. *International Journal of Automotive Technology*, 15 (6), 893-907.
- Nybacka, M., He, X., Su, Z., Drugge, L. & Bakker, E. (2014). Links between subjective assessments and objective metrics for steering, and evaluation of driver ratings. *Vehicle System Dynamics*, 52 (sup1), 31-50.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. *Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc*, 32 (3), 396-402.
- Oliver, R. L. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, 17 (4), 460.
- O'Muirheartaigh, C., Krosnick, J. A. & Helic, A. (2000). *Middle Alternatives, Acquiescence, and the Quality of Questionnaire Data*.
- Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, 49 (3), 197-237.
- Osgood, C. E. (1964). Semantic Differential Technique in the Comparative Study of Cultures. *American Anthropologist*, 66 (3), 171-200.
- ISO, 3888-1:1999 (1999). *Passenger cars - Test track for a severe lane-change manoeuvre - Part 1: Double lane-change*.
- ISO, 3888-2:2011 (2011). *Passenger cars - Test track for a severe lane-change manoeuvre - Part 2: Obstacle avoidance*.
- Pfeffer, P. E. & Scholz, H. (2010). Present-Day Cars - Subjective Evaluation of Steering Feel. In ATZ live (Hrsg.), *1st International Munich Chassis Symposium. chassis.tech plus* (Bd. 2010). Wiesbaden: Springer Vieweg.
- Pietsch, R. & Heißing, B. (2010). Modellbasierte Beurteilung des Lenkgefühls. In K. Becker (Hrsg.), *Subjektive Fahreindrücke sichtbar machen IV* (Haus der Technik Fachbuch, Bd. 108, S. 247-263). Renningen: expert-Verl.
- Pischinger, S. & Seiffert, U. (Hrsg.). (2016). *Vieweg Handbuch Kraftfahrzeugtechnik* (ATZ / MTZ-Fachbuch, 8., aktualisierte und erweiterte Auflage). Wiesbaden: Springer Vieweg.
- Porst, R. (2011). *Fragebogen. Ein Arbeitsbuch* (Studienskripten zur Soziologie, 3. Aufl.). Wiesbaden: VS Verl. für Sozialwiss.
- Preston, C. C. & Colman, A. M. (2000). Optimal number of response categories in rating scales. Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104 (1), 1-15.
- Renner, K.-H., Heydasch, T. & Ströhlein, G. (2012). *Forschungsmethoden der Psychologie. Von der Fragestellung zur Präsentation* (Basiswissen Psychologie). Wiesbaden: Springer VS.

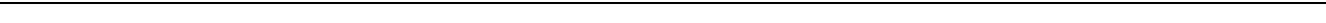
- Resch, S. & Mast, P. (2006). Engineered Emotion. In K. Becker (Hrsg.), *Subjektive Fahreindrücke sichtbar machen III. Korrelation zwischen objektiver Messung und subjektiver Beurteilung von Versuchsfahrzeugen und -komponenten* (Bd. 56, S. 118-126). Renningen: expert-Verl.
- Richman, W. L., Kiesler, S., Weisband, S. & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*, 84 (5), 754-775.
- Richter, M. & Flückiger, M. D. (2013). *Usability Engineering kompakt. Benutzbare Produkte gezielt entwickeln* (IT kompakt, 3. Aufl. 2013). Berlin: Springer.
- Riedel, A. & Arbinger, R. (1997). Subjektive und objektive Beurteilung des Fahrverhaltens von Pkw. *FAT - Schriftenreihe* (139).
- ISO, 13674-1:2010 (2010). *Road vehicles - Test method for the quantification of on-centre handling - Part 1: Weave test*.
- ISO, 13674-1:2010 (2010). *Road vehicles - Test method for the quantification of on-centre handling - Part 1: Weave test*. Schweiz.
- ISO/TR, 8725:1988 (1988). *Road vehicles - Transient open-loop response test method with one period of sinusoidal input*.
- Robert Bosch GmbH (Hrsg.). (2016). *Lenksysteme - Servotwin®*. Verfügbar unter http://www.bosch-automotive-steering.com/fileadmin/downloads/Flyer_Nkw/Servotwin_DE_lowres_160915.pdf
- Rocereto, J. F., Puzakova, M., Anderson, R. E. & Kwak, H. (2011). The Role of Response Formats on Extreme Response Style: A Case of Likert-Type vs. Semantic Differential Scales. In M. Sarstedt, M. Schwaiger & C. R. Taylor (Hrsg.), *Measurement and Research Methods in International Marketing* (Advances in International Marketing, Bd. 22, S. 53-71). Emerald Group Publishing Limited.
- Rossi, P. E., Gilula, Z. & Allenby, G. M. (2001). Overcoming Scale Usage Heterogeneity. A Bayesian Hierarchical Approach. *Journal of the American Statistical Association*, 96 (453), 20-31.
- Rothhämel, M. (2013). *Characterisation and Utilisation of Steering Feel in Heavy Trucks*. Dissertation, KTH Royal Institute of Technology. Sweden.
- Rothhämel, M., Ijkema, J. & Drugge, L. (2010). On a Method for Generating a Word Pool for the Description of Steering Feel. In *AVEC 10, the 10th International Symposium on Advanced Vehicle Control*. Loughborough: Loughborough Univ. Dep.
- Rothhämel, M., Ijkema, J. & Drugge, L. (2011a). A method to find correlations between steering feel and vehicle handling properties using a moving base driving simulator. *Vehicle System Dynamics*, 49 (12), 1837-1854.
- Rothhämel, M., Ijkema, J. & Drugge, L. (2011b). Finding Correlation between Steering Feel Assessments and the Drivers' Performance using a Moving Base Driving Simulator. In *Proceedings of the FAST-zero'11 1st International Symposium on Future Active Safety Technology towards zero traffic accidents*.
- Ruxton, G. D. & Neuhäuser, M. (2010). When should we use one-tailed hypothesis testing? *Methods in Ecology and Evolution*, 1 (2), 114-117.
- Saaty, T. L. & Ozdemir, M. S. (2003). Why the magic number seven plus or minus two. *Mathematical and Computer Modelling*, 38 (3-4), 233-244.
- Schalz, J., Duhr, A. & Marusic, Z. (2002). Subjektiv-objektiv Korrelation fahrdynamischer Größen in der Praxis. In K. Becker (Hrsg.), *Subjektive Fahreindrücke sichtbar machen II* (Haus der Technik Fachbuch, Bd. 12, S. 51-65). Renningen: expert-Verl.

- Schimmack, U. (2001). Pleasure, displeasure, and mixed feelings. Are semantic opposites mutually exclusive? *Cognition & Emotion*, 15 (1), 81-97.
- Schimmel, C. (2010). *Entwicklung eines fahrerbasierten Werkzeugs zur Objektivierung subjektiver Fahreindrücke*. Dissertation, Technische Universität München. München.
- Schimmel, C., Wijts, W., Jablonowski, C. & Agostini, J. (2014). Universal steering and suspension application with objective, subjective and virtual methods. In P. E. Pfeffer (Hrsg.), *5th International Munich Chassis Symposium 2014* (Proceedings, S. 91-106). Wiesbaden: Springer Fachmedien.
- Schmidt-Atzert, L. & Amelang, M. (2012). *Psychologische Diagnostik. Mit 82 Tabellen* (Springer-Lehrbuch, 5., vollständig überarbeitete und erweiterte Auflage). Berlin: Springer.
- Schmidtke, H. (Hrsg.). (1993). *Ergonomie* (3., neubearb. und erw. Aufl.). München: Hanser.
- Schoegg, P. & Ramschak, E. (2000). Vehicle Driveability Assessment using Neural Networks for Development, Calibration and Quality Tests. In *SAE 2000 World Congress* (SAE Technical Paper Series). SAE International.
- Schöttle, M. (2014). Nutzfahrzeugtrends Effizienz im Lkw und Bus. *ATZelektronik*, 9 (4), 62-65.
- Sedlmeier, P. & Renkewitz, F. (2011). *Forschungsmethoden und Statistik in der Psychologie* (PS Psychologie). München: Pearson Studium.
- Shyrokau, B., Loof, J., Stroosma, O., Wang, M. & Happee, R. (2015). Effect of steering model fidelity on subjective evaluation of truck steering feel. In H. Bülthoff, A. Kemeny & P. Pretto (Hrsg.), *Proceedings of DSC 2015 Europe Driving Simulation Conference & Exhibition. Tübingen, Germany, September 16 - 18, 2015* (S. 39-46). Tübingen.
- Striegel, K. (2016). *The Effect of Different Steering Torque Configurations on the Subjective Evaluation of Steering Feel in Heavy Trucks*. Bachelorarbeit, Universität Ulm. Ulm.
- Sturm, H.-J. (2011). *Markenfit und Markenwirkung*. Wiesbaden: Gabler.
- Tabachnick, B. G. & Fidell, L. S. (2014). *Using multivariate statistics* (Pearson custom library, 6. edition, Pearson new international edition). Harlow, Essex: Pearson Education.
- Tagesson, K. (2014). *Truck Steering System and Driver Interaction*. Thesis for licentiate of engineering, Chalmers University of Technology. Sweden.
- Thorndike, E. L. (1920). A constant error in psychological rating. *Journal of Applied Psychology*, 4 (1), 25-29.
- Tomaske, W. & Meywerk, M. (2006). Möglichkeiten zur Vermittlung von subjektiven Fahreindrücken mit Fahrsituationen. In K. Becker (Hrsg.), *Subjektive Fahreindrücke sichtbar machen III. Korrelation zwischen objektiver Messung und subjektiver Beurteilung von Versuchsfahrzeugen und -komponenten* (Bd. 56, S. 1-16). Renningen: expert-Verl.
- Turner, C. F. & Martin, E. (Eds.). (1984). *Surveying subjective phenomena. Volume 2*. New York: Russell Sage Foundation.
- Ullmann, S. (2006, April). *Der Normalfahrer als Messgröße für die Optimierung und Absicherung aktiver fahrdynamischer Regelsysteme*. 2. Tagung "Aktive Sicherheit durch Fahrerassistenz", Garching.
- Van der Laan, J. D., Heino, A. & de Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies*, 5 (1), 1-10.
- Weathers, D., Sharma, S. & Niedrich, R. W. (2005). The impact of the number of scale points, dispositional factors, and the status quo decision heuristic on scale reliability and response accuracy. *Journal of Business Research*, 58 (11), 1516-1524.
- Weber, E. H. (2012). *Der Tastsinn und das Gemeingefühl*. Hamburg: Tredition Classics.

- Weijters, B., Cabooter, E. & Schillewaert, N. (2010). The effect of rating scale format on response styles. The number of response categories and response category labels. *International Journal of Research in Marketing*, 27 (3), 236-247.
- Weinmann, O., Bitzer, F., Boos, N. & Burkhart, M. (2015). Lang-Lkw per Fernbedienung rangieren. In W. Siebenpfeiffer (Hrsg.), *Fahrerassistenzsysteme und Effiziente Antriebe* (ATZ/MTZ-Fachbuch). Wiesbaden: Springer Fachmedien.
- Weng, L.-J. (2004). Impact of the Number of Response Categories and Anchor Labels on Coefficient Alpha and Test-Retest Reliability. *Educational and Psychological Measurement*, 64 (6), 956-972.
- Wiesel, U. (2010). *Hybrides Lenksystem zur Kraftstoffeinsparung im schweren Nutzfahrzeug. Technische und methodische Ansätze* (Karlsruher Schriftenreihe Fahrzeugsystemtechnik, Bd. 1). Karlsruhe: KIT Scientific Publishing.
- Wiesel, U., Schwarzhaupt, A., Frey, M. & Gauterin, F. (2010). Hybride Lenkung zur Verbrauchs-Reduktion beim Nutzfahrzeug. *ATZ - Automobiltechnische Zeitschrift*, 112 (1), 62-68.
- Wolf, H. J. (2009). *Ergonomische Untersuchung des Lenkgefühls an Personenkraftwagen*. Dissertation, Technische Universität München. München.
- Yekutieli, D. & Benjamini, Y. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29 (4), 1165-1188.
- Zeitzen, F. (2012). *Sechs Kabinen auf einen Blick*, Eurotransport.de. Zugriff am 28.09.2016. Verfügbar unter <http://www.eurotransport.de/test/vergleich-sechs-fahrerhaeuser-auf-einen-blick-504794.html>
- Zeitzen, F. (2013). *Roter Lack gegen Rote Zahlen*, Eurotransport.de. Zugriff am 28.09.2016. Verfügbar unter <http://www.eurotransport.de/test/mercedes-actros-1842-ls-stream-space-roter-lack-gegen-rote-zahlen-5868546.html>
- ZF Friedrichshafen AG. (2014). *Zukunftsstudie 2014. Kapitel C: Berufskraftfahrerbefragung*. Zugriff am 28.09.2016. Verfügbar unter <http://www.zf-zukunftsstudie.de/wp-content/uploads/2014/08/ZF-Zukunftsstudie-2014-Kapitel-C.pdf>
- Zöller, I. M. (2015). *Analyse des Einflusses ausgewählter Gestaltungsparameter einer Fahrsimulation auf die Fahrerhaltensvalidität*. Dissertation, Technische Universität Darmstadt. Darmstadt.
- Zschocke, A. K. (2009). *Ein Beitrag zur objektiven und subjektiven Evaluierung des Lenkkomforts von Kraftfahrzeugen*. Dissertation, Universität Karlsruhe. Karlsruhe.
- Zschocke, A. K. & Albers, A. (2008). Links between subjective and objective evaluations regarding the steering character of automobiles. *International Journal of Automotive Technology*, 9 (4), 473-481.
- Zwick, W. R. & Velicer, W. F. (1984). *A Comparison of Five Rules for Determining the Number of Components in Data Sets*. Annual Meeting of the American Psychological Association.



8. Anhang



Anhang I Studie I: Deskriptive Statistiken zu Hypothese 1

Variable	Aspekt	<i>N</i>	<i>Mdn</i>	<i>IQR</i>
Leichtgängig (Rangieren) – Niveau	F_HRST1 – FZG1	27	0,013	1,738
	F_HRST1 – FZG2	27	-0,422	2,173
	F_HRST2 – FZG1	6	0,882	1,521
	F_HRST2 – FZG2	6	0,665	1,304
Leichtgängig (Rangieren) – Gefallen	F_HRST1 – FZG1	27	0,287	1,502
	F_HRST1 – FZG2	27	-0,464	2,253
	F_HRST2 – FZG1	6	0,475	1,784
	F_HRST2 – FZG2	6	0,475	0,939
Lenkungsstößigkeit – Niveau	F_HRST1 – FZG1	27	0,858	0,425
	F_HRST1 – FZG2	27	-0,416	1,699
	F_HRST2 – FZG1	6	0,645	0,849
	F_HRST2 – FZG2	6	0,008	0,637
Lenkungsstößigkeit – Gefallen	F_HRST1 – FZG1	27	0,817	0,451
	F_HRST1 – FZG2	27	-0,535	1,802
	F_HRST2 – FZG1	6	0,592	0,901
	F_HRST2 – FZG2	6	0,141	1,352
Agil – Niveau	F_HRST1 – FZG1	26	0,593	0,544
	F_HRST1 – FZG2	26	0,593	0,680
	F_HRST2 – FZG1	6	0,049	1,497
	F_HRST2 – FZG2	6	0,865	1,497
Agil – Gefallen	F_HRST1 – FZG1	26	0,488	0,542
	F_HRST1 – FZG2	26	0,217	1,084
	F_HRST2 – FZG1	6	0,217	1,762
	F_HRST2 – FZG2	6	0,759	1,355
Direkt – Niveau	F_HRST1 – FZG1	27	0,414	1,407
	F_HRST1 – FZG2	27	0,062	1,407
	F_HRST2 – FZG1	6	0,414	2,110

Direkt – Gefallen	F_HRST2 – FZG2	6	0,941	1,582
	F_HRST1 – FZG1	27	0,502	0,954
	F_HRST1 – FZG2	27	0,184	1,590
	F_HRST2 – FZG1	6	0,184	1,749
Hat Spiel – Niveau	F_HRST2 – FZG2	6	0,820	1,033
	F_HRST1 – FZG1	27	0,562	0,955
	F_HRST1 – FZG2	27	0,084	1,433
	F_HRST2 – FZG1	6	-0,155	1,075
Hat Spiel – Gefallen	F_HRST2 – FZG2	6	0,562	0,597
	F_HRST1 – FZG1	27	0,574	0,887
	F_HRST1 – FZG2	27	0,574	1,331
	F_HRST2 – FZG1	6	-0,535	2,440
Gesamt-Gefallenbewertung	F_HRST2 – FZG2	6	0,352	1,775
	F_HRST1 – FZG1	27	0,777	0,734
	F_HRST1 – FZG2	27	0,043	0,734
	F_HRST2 – FZG1	6	0,043	1,834
	F_HRST2 – FZG2	6	0,410	1,284
Variable	Sub-Stichprobe	<i>N</i>	<i>M</i>	<i>SD</i>
FZG1 - Niveau	F_HRST1	269	0,344	0,757
	F_HRST2	60	0,115	1,072
FZG1 - Gefallen	F_HRST1	269	0,422	0,731
	F_HRST2	60	0,068	1,002
FZG2 - Niveau	F_HRST1	269	-0,096	1,058
	F_HRST2	60	0,350	0,826
FZG2 - Gefallen	F_HRST1	269	-0,208	1,085
	F_HRST2	60	0,345	0,714

Anhang II Studie I: Deskriptive Statistiken zu Hypothese 2

Variable	Sub-Stichprobe	<i>N</i>	<i>M</i>	<i>SD</i>
FZG1 - Niveau	Normalfahrer	409	0,256	0,869
	Experten	100	-0,626	1,064
FZG1 - Gefallen	Normalfahrer	450	0,350	0,774
	Experten	110	-0,729	1,133
FZG2 - Niveau	Normalfahrer	409	-0,080	1,030
	Experten	100	-0,093	0,927
FZG2 - Gefallen	Normalfahrer	450	-0,150	1,037
	Experten	110	-0,089	0,922



Anhang III Studie I: Deskriptive Statistiken zu Hypothese 3

Variable	Sub-Stichprobe	<i>N</i>	<i>M</i>	<i>SD</i>
Effektstärke <i>r</i>	Normalfahrer	21	0,369	0,199
	Experten	21	0,632	0,216
Standardabweichung <i>SD</i>	Normalfahrer	21	1,916	0,438
	Experten	21	2,022	0,412

Anhang IV

Studie I: Deskriptive Statistiken zur explorativen Analyse

Variable	Sub-Stichprobe	<i>N</i>	<i>M</i>	<i>SD</i>
Gesamt-Gefallenbewertung FZG1	Normalfahrer	41	0,401	0,789
	Experten	10	-0,911	0,981
Gesamt-Gefallenbewertung FZG2	Normalfahrer	41	-0,261	1,075
	Experten	10	0,337	0,513

Anhang V Studie II: Deskriptive Statistiken zu Hypothese 1a

Variable	<i>N</i>	<i>Mdn</i>	<i>IQR</i>
Leichtgängig - Niveau - FZG1	10	-0,233	1,492
Leichtgängig - Niveau - FZG2	10	1,110	1,045
Leichtgängig - Niveau - FZG3	10	-0,382	1,343
Leichtgängig - Niveau - FZG4	10	-0,382	1,865
Leichtgängig - Gefallen - FZG1	10	-0,078	1,577
Leichtgängig - Gefallen - FZG2	10	1,123	1,277
Leichtgängig - Gefallen - FZG3	10	-0,228	1,352
Leichtgängig - Gefallen - FZG4	10	-0,529	1,952
Haltemoment - Niveau - FZG1	10	-0,409	1,947
Haltemoment - Niveau - FZG2	10	-0,409	1,095
Haltemoment - Niveau - FZG3	10	-0,166	1,339
Haltemoment - Niveau - FZG4	10	-0,409	0,487
Haltemoment - Gefallen - FZG1	10	0,155	0,746
Haltemoment - Gefallen - FZG2	10	0,155	1,044
Haltemoment - Gefallen - FZG3	10	0,155	0,149
Haltemoment - Gefallen - FZG4	10	0,155	0,597
Rückstellverhalten Geschwindigkeit - Niveau - FZG1	10	0,258	1,450
Rückstellverhalten Geschwindigkeit - Niveau - FZG2	10	1,063	0,725
Rückstellverhalten Geschwindigkeit - Niveau - FZG3	10	-0,387	2,014
Rückstellverhalten Geschwindigkeit - Niveau - FZG4	10	-1,192	0,725
Rückstellverhalten Geschwindigkeit - Gefallen - FZG1	10	0,388	1,479
Rückstellverhalten Geschwindigkeit - Gefallen - FZG2	10	0,881	1,315
Rückstellverhalten Geschwindigkeit - Gefallen - FZG3	10	0,059	2,301
Rückstellverhalten Geschwindigkeit - Gefallen - FZG4	10	-0,270	2,055
Rückstellverhalten Restlenkwinkel - Niveau - FZG1	10	+0,685	1,511
Rückstellverhalten Restlenkwinkel - Niveau - FZG2	10	+1,021	1,427
Rückstellverhalten Restlenkwinkel - Niveau - FZG3	10	-0,490	1,175

Rückstellverhalten Restlenkwinkel - Niveau - FZG4	10	-0,322	1,343
Rückstellverhalten Restlenkwinkel - Gefallen - FZG1	10	0,671	1,385
Rückstellverhalten Restlenkwinkel - Gefallen - FZG2	10	0,962	0,437
Rückstellverhalten Restlenkwinkel - Gefallen - FZG3	10	-0,350	2,041
Rückstellverhalten Restlenkwinkel - Gefallen - FZG4	10	-0,350	1,239

Anhang VI Studie II: Deskriptive Statistiken zu Hypothese 1b

Variable	<i>N</i>	<i>Mdn</i>	<i>IQR</i>
Lenkungsstößigkeit FZG1 - Niveau	10	+0,200	1,941
Lenkungsstößigkeit FZG2 - Niveau	10	+1,124	1,664
Lenkungsstößigkeit FZG3 - Niveau	10	-0,170	1,017
Lenkungsstößigkeit FZG4 - Niveau	10	-0,355	1,202
Lenkungsstößigkeit FZG1 - Gefallen	10	-0,344	1,814
Lenkungsstößigkeit FZG2 - Gefallen	10	0,802	1,432
Lenkungsstößigkeit FZG3 - Gefallen	10	0,229	1,814
Lenkungsstößigkeit FZG4 - Gefallen	10	-0,535	0,955



Anhang VII Studie II: Deskriptive Statistiken zu Hypothese 1c

Variable	<i>N</i>	<i>Mdn</i>	<i>IQR</i>
Gesamt-Gefallenbewertung FZG1	10	-0,308	0,290
Gesamt-Gefallenbewertung FZG2	10	-0,308	1,452
Gesamt-Gefallenbewertung FZG3	10	-0,308	1,162
Gesamt-Gefallenbewertung FZG4	10	0,854	2,324

Anhang VIII Studie II: Deskriptive Statistiken zu Hypothese 2a, 2b und 2c

Variable	<i>N</i>	<i>M</i>	<i>SD</i>
Bewertungen Hinfahrt - Niveau	1.287	0,000	0,991
Bewertungen Rückfahrt - Niveau	1.287	0,002	0,981
Bewertungen Hinfahrt - Gefallen	1.403	0,002	0,991
Bewertungen Rückfahrt - Gefallen	1.403	-0,002	0,991

Variable	<i>N</i>	<i>M</i>	<i>SD</i>
Reguläre Fahrt - Niveau	290	-0,082	1,026
Wiederholungsfahrt - Niveau	290	0,082	0,916
Reguläre Fahrt - Gefallen	290	-0,092	1,016
Wiederholungsfahrt - Gefallen	290	0,092	0,926

Variable	<i>N</i>	<i>M</i>	<i>SD</i>
Differenzen zwischen regulären Fahrten - Niveau	289	0,988	0,615
Differenzen zwischen regulärer Fahrt & Wiederholungsfahrt - Niveau	289	0,075	1,032
Differenzen zwischen regulären Fahrten - Gefallen	290	0,992	0,619
Differenzen zwischen regulärer Fahrt & Wiederholungsfahrt - Gefallen	290	-0,208	0,887

Anhang IX Studie II: Deskriptive Statistiken und statistische Kennwerte zur explorativen Analyse

Variable	<i>N</i>	<i>Mdn</i>	<i>IQR</i>
Lenkaktivität - reguläre Fahrt - Gefallen	10	-0,69	1,72
Lenkaktivität Wiederholungsfahrt - Gefallen	10	0,12	1,47
Mittenzentrierung - reguläre Fahrt - Gefallen	10	-0,10	1,32
Mittenzentrierung - Wiederholungsfahrt - Gefallen	10	0,41	0,90
Korrekturbedarf Häufigkeit (Querneigung) - reguläre Fahrt - Gefallen	10	-0,15	0,86
Korrekturbedarf Häufigkeit (Querneigung) - Wiederholungsfahrt - Gefallen	10	0,41	0,54

Variable	<i>N</i>	<i>z</i>	<i>p</i> _{unkorrigiert}	<i>p</i> _{korrigiert}
Mittenzentrierung	10	-2,55	.011	.315
Lenkaktivität	10	-2,55	.011	.158
Korrekturbedarf Querneigung Häufigkeit	10	-2,11	.035	.336
agil	10	-1,81	.071	.513
komfortabel	10	-1,73	.084	.485
genau	10	-1,63	.102	.495
flattert	10	-1,47	.143	.591
direkt	10	-1,44	.151	.547
Korrekturbedarf - Lenkwinkel (Querneigung)	10	-1,38	.168	.540
Spiel	10	-1,38	.168	.488
Anlenkmoment	10	-1,34	.180	.474
verzögert	10	-1,19	.234	.566
Fahrbahnkontakt	10	-1,08	.281	.627
knarzt	10	-0,94	.348	.721
nervös	10	-0,85	.394	.761
synthetisch	10	-0,85	.396	.718
sicher	10	-0,71	.480	.818
Lenkungsstößigkeit	10	-0,57	.570	.919

Rückstellverhalten - Geschwindigkeit	10	-0,42	.673	1.027
Korrekturbedarf Häufigkeit (Geradeausfahrt)	10	-0,33	.739	1.071
hakt	10	-0,32	.748	1.033
gleichmäßig	10	-0,25	.799	1.053
Rückstellverhalten - Restlenkwinkel	10	-0,18	.857	1.081
Korrekturbedarf Lenkwinkel (Geradeausfahrt)	10	-0,18	.861	1.040
Haltemoment	10	-0,17	.863	1.001
leichtgängig	10	-0,17	.865	.964
Korrekturbedarf - Häufigkeit (Kurvenfahrt)	10	0,00	1.000	1.074
Korrekturbedarf Lenkwinkel (Kurvenfahrt)	10	0,00	1.000	1.036
vibriert	10	0,00	1.000	1.000

Anhang X Studie III: Statistische Kennwerte zu Hypothese 2

Variable	Komponente	<i>B</i>	Standard- fehler <i>B</i>	β	<i>t</i>	<i>p</i>
Lenkmoment (Rangieren) - Gefallen	Konstante	,099	,047		2,129	.034
	Lenkmoment (Rangieren) Niveau	-,751	,031	-0,751	-24,316	< .001
	Lenkmoment (Rangieren) Niveau ²	-,100	,035	-0,087	-2,821	.005
Lenkmoment (Slalom) - Gefallen	Konstante	,506	,067		7,557	< .001
	Lenkmoment (Slalom) Niveau	,067	,044	0,067	1,525	.128
	Lenkmoment (Slalom) Niveau ²	-,507	,051	-0,431	-9,864	< .001
Anlenkmoment - Gefallen	Konstante	,415	,065		6,436	< .001
	Anlenkmoment Niveau	-,112	,044	-0,112	-2,553	.011
	Anlenkmoment Niveau ²	-,125	,014	-0,383	-8,725	< .001
Haltemoment - Gefallen	Konstante	,240	,055		4,341	< .001
	Haltemoment Niveau	-,585	,037	-0,585	-15,607	< .001
	Haltemoment Niveau ²	-,240	,041	-0,221	-5,896	< .001
Rückstellverhalten -Gefallen	Konstante	-1,113E- 16	,019		,000	1.000
	Rückstellverhalten Niveau	,918	,019	0,918	48,971	< .001
Lenkpräzision - Gefallen	Konstante	,001	,019		,077	.939
	Lenkpräzision Niveau	,911	,019	0,912	46,964	< .001
Fahrbahnkontakt - Gefallen	Konstante	,067	,029		2,284	.023
	Fahrbahnkontakt Niveau	,882	,022	0,883	39,816	< .001
	Fahrbahnkontakt Niveau ²	-,065	,019	-0,074	-3,320	.001
Mittenzentrierung - Gefallen	Konstante	1,493E- 16	,032		,000	1.000
	Mittenzentrierung Niveau	,732	,032	0,732	22,594	< .001

Anhang XI Studie III: Statistische Kennwerte zu Hypothese 3a

Variante 1							
Vergleich	Variable	<i>df</i>	<i>t</i>	<i>p</i>	<i>N</i>	<i>M</i>	<i>SD</i>
1	Haltemoment - Niveau	17	1,68	.217	18	-1,299	0,379
	Haltemoment - Niveau weit					-1,478	0,230
2	Haltemoment - Gefallen	17	1,68	.408	18	0,721	1,010
	Haltemoment - Gefallen getrennt					0,387	1,339
3	Lenkmoment (Rangieren) - Niveau	19	-0,81	.556	20	-1,218	0,254
	Lenkmoment (Rangieren) - Niveau unmittelbar					-1,111	0,652
4	Lenkmoment (Rangieren) - Gefallen	19	-0,81	.662	20	0,716	0,996
	Lenkmoment (Rangieren) - Gefallen unmittelbar					0,883	1,016
5	Lenkmoment (Slalom) - Niveau	35	1,04	.434	36	-1,040	0,268
	Lenkmoment (Slalom) - Niveau unmittelbar					-1,116	0,360
6	Lenkmoment (Slalom) - Niveau	37	0,00	1.000	38	-1,058	0,272
	Lenkmoment (Slalom) - Niveau getrennt					-1,058	0,488
7	Lenkmoment (Slalom) - Gefallen	35	0,90	.497	36	-0,451	1,104
	Lenkmoment (Slalom) - Gefallen unmittelbar					-0,602	1,130
8	Lenkmoment (Slalom) - Gefallen	37	1,25	.344	38	-0,081	1,180
	Lenkmoment (Slalom) - Gefallen getrennt					-0,271	1,280
9	Rückstellverhalten - Niveau	18	-0,75	.581	19	0,090	1,209
	Rückstellverhalten - Niveau unmittelbar					0,216	1,013
10	Rückstellverhalten - Niveau	17	0,00	1.000	18	-0,317	1,078
	Rückstellverhalten - Niveau getrennt					-0,317	1,209
11	Rückstellverhalten - Gefallen	18	-1,28	.345	19	0,123	1,242
	Rückstellverhalten - Gefallen unmittelbar					0,348	0,988
12	Rückstellverhalten - Gefallen	17	-0,26	.867	18	-0,594	0,888
	Rückstellverhalten - Gefallen getrennt					-0,526	1,058

Variante 2							
Vergleich	Variable	<i>df</i>	<i>t</i>	<i>p</i>	<i>N</i>	<i>M</i>	<i>SD</i>
1	Lenkpräzision - Niveau	37	0,16	.919	38	0,206	1,102
	Lenkpräzision - Niveau unmittelbar					0,173	0,975
2	Lenkpräzision - Niveau	17	-1,61	.234	18	-0,235	0,554
	Lenkpräzision - Niveau getrennt					0,159	1,036
3	Lenkpräzision - Gefallen	37	0,08	.973	38	0,162	1,017
	Lenkpräzision - Gefallen unmittelbar					0,145	0,955
4	Lenkpräzision - Gefallen	17	-1,13	.404	18	-0,192	0,627
	Lenkpräzision - Gefallen getrennt					0,056	1,032
5	Mittenzentrierung - Niveau	36	-1,45	.279	37	-0,312	0,763
	Mittenzentrierung - Niveau unmittelbar					-0,081	0,842
6	Mittenzentrierung - Niveau	36	-1,94	.137	37	-0,276	0,610
	Mittenzentrierung - Niveau getrennt					0,131	0,989
7	Mittenzentrierung - Gefallen	36	-1,01	.441	37	-0,237	0,812
	Mittenzentrierung - Gefallen unmittelbar					-0,067	0,841
8	Mittenzentrierung - Gefallen	36	-3,17	.008	37	-0,293	0,616
	Mittenzentrierung - Gefallen getrennt					0,328	0,945

Variante 3

Vergleich	Variable	<i>df</i>	<i>t</i>	<i>p</i>	<i>N</i>	<i>M</i>	<i>SD</i>
1	Fahrbahnkontakt - Niveau	17	1,37	.323	18	0,200	0,593
	Fahrbahnkontakt - Niveau getrennt					-0,009	0,712
2	Fahrbahnkontakt - Gefallen	17	1,38	.325	18	0,307	0,616
	Fahrbahnkontakt - Gefallen getrennt					0,053	0,752
3	Haltemoment - Niveau	19	,81	.549	20	0,042	0,444
	Haltemoment - Niveau unmittelbar					-0,065	0,437
4	Haltemoment - Gefallen	19	-2,90	.023	20	0,120	0,390
	Haltemoment - Gefallen unmittelbar					0,541	0,588
5	Lenkmoment (Rangieren) - Niveau	18	1,32	.332	19	0,192	0,649
	Lenkmoment (Rangieren) - Niveau unmittelbar					-0,035	0,565
6	Lenkmoment (Rangieren) - Niveau	34	-1,03	.435	35	-0,194	0,542
	Lenkmoment (Rangieren) - Niveau getrennt					-0,040	0,596
7	Lenkmoment (Rangieren) - Gefallen	18	-1,02	.439	19	-0,014	0,596
	Lenkmoment (Rangieren) - Gefallen unmittelbar					0,192	0,710
8	Lenkmoment (Rangieren) - Gefallen	34	0,20	.910	35	0,202	0,526
	Lenkmoment (Rangieren) - Gefallen getrennt					0,170	0,709

Variante 4

Vergleich	Variable	<i>df</i>	<i>t</i>	<i>p</i>	<i>N</i>	<i>M</i>	<i>SD</i>
1	Anlenkmoment - Niveau	37	0,32	.834	38	0,784	0,397
	Anlenkmoment - Niveau unmittelbar					0,755	0,438
2	Anlenkmoment - Niveau	36	-0,18	.910	37	0,495	0,288
	Anlenkmoment - Niveau getrennt					0,510	0,474
3	Anlenkmoment - Gefallen	37	-0,90	.504	38	0,041	1,040
	Anlenkmoment - Gefallen unmittelbar					0,146	0,951
4	Anlenkmoment - Gefallen	36	-1,20	.361	37	-0,039	0,649
	Anlenkmoment - Gefallen getrennt					0,122	0,769
5	Rückstellverhalten - Niveau	17	-0,41	.789	18	-0,083	0,832
	Rückstellverhalten - Niveau getrennt					0,017	1,089
6	Rückstellverhalten - Gefallen	17	0,00	1.000	18	-0,018	0,891
	Rückstellverhalten - Gefallen getrennt					-0,018	1,070

Variante 5

Vergleich	Variable	<i>df</i>	<i>t</i>	<i>p</i>	<i>N</i>	<i>M</i>	<i>SD</i>
1	Fahrbahnkontakt - Niveau	35	1,58	.233	36	0,637	0,801
	Fahrbahnkontakt - Niveau unmittelbar					0,393	0,957
2	Fahrbahnkontakt - Niveau	18	-0,38	.793	19	0,524	0,780
	Fahrbahnkontakt - Niveau getrennt					0,590	1,173
3	Fahrbahnkontakt - Gefallen	35	1,32	.326	36	0,434	0,803
	Fahrbahnkontakt - Gefallen unmittelbar					0,234	0,998
4	Fahrbahnkontakt - Gefallen	18	0,32	.830	19	0,506	0,667
	Fahrbahnkontakt - Gefallen getrennt					0,437	1,160
5	Haltemoment - Niveau	18	0,68	.625	19	1,204	0,583
	Haltemoment - Niveau unmittelbar					1,119	0,522
6	Haltemoment - Niveau	17	-1,32	.331	18	0,972	0,540
	Haltemoment - Niveau getrennt					1,121	0,390
7	Haltemoment - Gefallen	18	-1,14	.400	19	-1,145	0,486
	Haltemoment - Gefallen unmittelbar					-1,050	0,547
8	Haltemoment - Gefallen	17	0,00	1.000	18	-0,648	0,819
	Haltemoment - Gefallen getrennt					-0,648	0,917
9	Rückstellverhalten - Niveau	19	0,75	.582	20	-0,207	0,880
	Rückstellverhalten - Niveau unmittelbar					-0,357	0,995
10	Rückstellverhalten - Gefallen	19	0,59	.667	20	-0,055	0,980
	Rückstellverhalten - Gefallen unmittelbar					-0,177	1,015
11	Lenkpräzision - Niveau	18	1,78	.187	19	0,189	0,841
	Lenkpräzision - Niveau getrennt					-0,320	0,943
12	Lenkpräzision - Gefallen	17	1,96	.148	18	0,056	0,935
	Lenkpräzision - Gefallen getrennt					-0,547	1,000

Anhang XII Studie III: Statistische Kennwerte zu Hypothese 3b

Variable	<i>df</i>	<i>t</i>	<i>p</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Lenkmoment (Rangieren) Niveau V3	73	-0,64	.642	74	-0,081	1,095
Lenkmoment (Slalom) Niveau V3	73	0,40	.783	74	0,041	0,867
Anlenkmoment Niveau V3	74	-0,49	.728	75	-0,053	0,943
Haltemoment Niveau V3	74	1,72	.191	75	0,160	0,806
Rückstellverhalten Niveau V3	74	2,45	.041	75	0,333	1,178
Lenkpräzision Niveau V3	74	1,71	.191	75	0,200	1,013
Fahrbahnkontakt Niveau V3	73	1,56	.236	74	0,189	1,043
Mittenzentrierung Niveau V3	73	-1,51	.248	74	-0,189	1,081
Lenkmoment (Rangieren) Gefallen V3	73	0,82	.543	74	0,200	1,013
Lenkmoment (Slalom) Gefallen V3	73	1,73	.192	74	0,120	0,869
Anlenkmoment Gefallen V3	74	1,20	.363	75	-0,203	1,060
Haltemoment Gefallen V3	74	0,52	.712	75	-0,053	0,943
Rückstellverhalten Gefallen V3	74	2,21	.071	75	0,189	1,043
Lenkpräzision Gefallen V3	74	1,92	.136	75	0,293	1,148
Fahrbahnkontakt Gefallen V3	73	1,32	.326	74	0,227	1,021
Mittenzentrierung Gefallen V3	73	-1,65	.208	74	0,041	0,867

Anhang XIII Studie III: Deskriptive Statistiken zu Hypothese 4

Variable	Variante	<i>N</i>	<i>M</i>	<i>SD</i>
Lenkmoment (Rangieren) - Niveau	V1	74	-1,202	0,272
	V2	74	-0,612	0,538
	V3	74	-0,022	0,590
	V4	74	0,838	0,440
	V5	74	1,326	0,357
Lenkmoment (Slalom) - Niveau	V1	74	-1,049	0,269
	V2	74	-0,351	0,460
	V3	74	0,281	0,477
	V4	74	0,913	0,393
	V5	74	1,292	0,407
Anlenkmoment- Niveau	V1	75	-1,462	0,408
	V2	75	-0,795	0,501
	V3	75	-0,114	0,518
	V4	75	0,641	0,374
	V5	75	1,096	0,401
Haltemoment- Niveau	V1	75	-1,275	0,319
	V2	75	-0,709	0,455
	V3	75	0,102	0,433
	V4	75	0,711	0,521
	V5	75	0,991	0,550

Anhang XIV Studie III: Statistische Kennwerte zu Hypothese 5

Variable	Komponente	<i>B</i>	Standard- fehler <i>B</i>	β	<i>t</i>	<i>p</i>
Lenkmoment (Rangieren) - Niveau	Konstante	-1,239	,242		-5,120	< .001
	Lenkmoment	-,294	,169	-0,287	-1,746	.082
	Lenkmoment ²	,193	,027	1,161	7,070	< .001
Lenkmoment (Slalom) - Niveau	Konstante	-2,073	,200		-10,367	< .001
	Lenkmoment	,492	,147	0,535	3,345	.001
	Lenkmoment ²	,057	,024	0,379	2,368	.018
Anlenkmoment - Niveau	Konstante	-2,107	,254		-8,303	< .001
	Lenkmoment	,167	,176	0,161	,950	.343
	Lenkmoment ²	,122	,028	0,735	4,333	< .001
Haltemoment - Niveau	Konstante	-2,309	,235		-9,844	< .001
	Lenkmoment	,460	,163	0,473	2,826	.005
	Lenkmoment ²	,065	,026	0,418	2,495	.013
Rückstellverhalten - Niveau	Konstante	-2,778	,482		-5,762	< .001
	Lenkmoment	2,178	,344	2,319	6,332	< .001
	Lenkmoment ²	-,368	,056	-2,426	-6,626	< .001

Anhang XV Studie III: Statistische Kennwerte zu Hypothese 6

Variable	Komponente	<i>B</i>	Standard- fehler <i>B</i>	β	<i>t</i>	<i>p</i>
Lenkmoment (Rangieren) - Gefallen	Konstante	-0,078	,373		-0,209	.,835
	Lenkmoment	0,943	,260	0,918	3,626	< .001
	Lenkmoment ²	-0,261	,042	-1,571	-6,206	< .001
Lenkmoment (Slalom) - Gefallen	Konstante	-3,788	,453		-8,358	< .001
	Lenkmoment	2,810	,333	3,058	8,432	< .001
	Lenkmoment ²	-0,456	,055	-3,016	-8,317	< .001
Anlenkmoment - Gefallen	Konstante	-4,985	,513		-9,711	< .001
	Lenkmoment	3,568	,356	3,441	10,033	< .001
	Lenkmoment ²	-0,572	,057	-3,447	-10,050	< .001
Haltemoment - Gefallen	Konstante	-1,505	,429		-3,511	< .001
	Lenkmoment	1,719	,297	1,769	5,781	< .001
	Lenkmoment ²	-0,348	,047	-2,248	-7,348	< .001
Rückstellverhalten - Gefallen	Konstante	-2,908	,484		-6,012	< .001
	Lenkmoment	2,217	,345	2,361	6,426	< .001
	Lenkmoment ²	-0,368	,056	-2,425	-6,600	< .001

Anhang XVI Studie III: Deskriptive Statistiken zur explorativen Analyse

Variable	Gruppe	<i>N</i>	<i>M</i>	<i>SD</i>
Fahrerfahrung	Gering	380	0,029	1,046
	Mittel	352	0,151	0,950
	Hoch	430	0,133	1,037
Alter	Jünger	412	0,128	1,009
	Mittleres Alter	398	0,073	1,075
	Älter	382	0,060	0,963

Anhang XVII Studie III: Deskriptive Statistiken der Gestaltungsempfehlungen

Bewertungskriterium	Variante	<i>N</i>	<i>M</i>	<i>SD</i>
Lenkmoment (Rangieren) - Gefallen	V1	74	0,753	0,920
	V2	74	0,504	0,770
	V3	74	0,134	0,630
	V4	74	-0,642	0,595
	V5	74	-1,117	0,693
Lenkmoment (Slalom) - Gefallen	V1	74	-0,261	1,151
	V2	74	0,604	0,818
	V3	74	0,383	0,650
	V4	74	0,041	0,770
	V5	74	-0,334	0,836
Anlenkmoment - Gefallen	V1	75	-0,580	1,212
	V2	75	0,530	0,915
	V3	75	0,451	0,575
	V4	75	0,001	0,865
	V5	75	-0,536	0,895
Haltemoment - Gefallen	V1	75	0,368	1,254
	V2	75	0,769	0,690
	V3	75	0,144	0,400
	V4	75	-0,393	0,750
	V5	75	-0,705	0,783
Rückstellverhalten - Gefallen	V1	75	-0,244	1,130
	V2	75	0,342	0,925
	V3	75	0,399	0,701
	V4	75	0,073	0,896
	V5	75	-0,480	0,966
Lenkpräzision - Gefallen	V1	74	-0,821	1,113
	V2	74	0,102	0,939

	V3	74	0,308	0,656
	V4	74	0,196	0,796
	V5	74	0,265	0,985
Fahrbahnkontakt - Gefallen	V1	73	-1,074	0,902
	V2	73	-0,089	0,841
	V3	73	0,278	0,692
	V4	73	0,385	0,920
	V5	73	0,287	0,862
Mittenzentrierung - Gefallen	V1	74	-0,576	1,123
	V2	74	-0,265	0,716
	V3	74	0,036	0,738
	V4	74	0,253	0,941
	V5	74	0,422	1,167

Bewertungskriterium	Variante	<i>N</i>	<i>p</i>
Lenkmoment (Rangieren) - Gefallen	V1 vs. V2	74	.227
	V1 vs. V3	74	.004
	V1 vs. V4	74	< .000
	V1 vs. V5	74	< .000
	V2 vs. V3	74	.124
	V2 vs. V4	74	.007
	V2 vs. V5	74	.001
	V3 vs. V4	74	.683
	V3 vs. V5	74	.158
	V4 vs. V5	74	.964
Lenkmoment (Slalom) - Gefallen	V1 vs. V2	74	< .000
	V1 vs. V3	74	.001
	V1 vs. V4	74	.594

	V1 vs. V5	74	1.000
	V2 vs. V3	74	.531
	V2 vs. V4	74	< .000
	V2 vs. V5	74	< .000
	V3 vs. V4	74	.122
	V3 vs. V5	74	< .000
	V4 vs. V5	74	.005
Anlenkmoment - Gefallen	V1 vs. V2	75	< .000
	V1 vs. V3	75	< .000
	V1 vs. V4	75	.019
	V1 vs. V5	75	1.000
	V2 vs. V3	75	.999
	V2 vs. V4	75	.028
	V2 vs. V5	75	< .000
	V3 vs. V4	75	.007
	V3 vs. V5	75	< .000
	V4 vs. V5	75	< .000
Haltemoment - Gefallen	V1 vs. V2	75	.054
	V1 vs. V3	75	.787
	V1 vs. V4	75	.001
	V1 vs. V5	75	< .000
	V2 vs. V3	75	< .000
	V2 vs. V4	75	< .000
	V2 vs. V5	75	< .000
	V3 vs. V4	75	< .000
	V3 vs. V5	75	< .000
	V4 vs. V5	75	.007
Rückstellverhalten - Gefallen	V1 vs. V2	75	.001
	V1 vs. V3	75	.001
	V1 vs. V4	75	.542

	V1 vs. V5	75	.892
	V2 vs. V3	75	1.000
	V2 vs. V4	75	.589
	V2 vs. V5	75	< .000
	V3 vs. V4	75	.056
	V3 vs. V5	75	< .000
	V4 vs. V5	75	< .000
Lenkpräzision - Gefallen	V1 vs. V2	74	< .000
	V1 vs. V3	74	< .000
	V1 vs. V4	74	< .000
	V1 vs. V5	74	< .000
	V2 vs. V3	74	.588
	V2 vs. V4	74	1.000
	V2 vs. V5	74	.990
	V3 vs. V4	74	.993
	V3 vs. V5	74	1.000
	V4 vs. V5	74	1.000
Fahrbahnkontakt - Gefallen	V1 vs. V2	73	< .000
	V1 vs. V3	73	< .000
	V1 vs. V4	73	< .000
	V1 vs. V5	73	< .000
	V2 vs. V3	73	.019
	V2 vs. V4	73	.043
	V2 vs. V5	73	.210
	V3 vs. V4	73	.997
	V3 vs. V5	73	1.000
	V4 vs. V5	73	.992
Mittenzentrierung - Gefallen	V1 vs. V2	74	.227
	V1 vs. V3	74	.004
	V1 vs. V4	74	< .000

V1 vs. V5	74	< .000	
V2 vs. V3	74	.124	
V2 vs. V4	74	.007	
V2 vs. V5	74	.001	
V3 vs. V4	74	.683	
V3 vs. V5	74	.158	
V4 vs. V5	74	.964	